intento e2f

The State of Machine Translation 2023

An independent multi-domain evaluation of MT Engines

37 MT Engines

22 Language Pairs

9 Content Domains

Disclaimer

April 1—April 24, 2023

The MT systems used in this report were accessed from April 1 to April 24, 2023. Some of these systems may have changed since then.

Automatic scoring

This report demonstrates the performance of those systems exclusively on the datasets used for this report (see slide 12) using semantic similarity scores. The final MT decision requires Human LQA and depends on each specific use case.

Stock models only

If you consider customizing NMT, your choice may vary from what is suggested here. In the solutions we build for our clients, we often see Amazon, Globalese, Google, Microsoft, ModernMT, and Systran being top choices, depending on languages and the training data volume.

Plain text only

The evaluation was done on plain text data. We often see different results for tagged text (like those found in CAT/TMS systems) for some MT vendors and language pairs due to imperfect inline tag support.

Valid for a specific dataset

We run multiple evaluations for our clients using various language pairs and domains, and often observe different MT system rankings than those provided in this report.

There's no "best" MT system

MT performance depends on how similar your data is to the data used to train the vendors' models, their algorithms, and your quality requirements.

Trademarks

All third-party trademarks, registered trademarks, product names, and company names or logos mentioned in the Report are the property of their respective owners, and the use of such Third-Party Trademarks inures to the benefit of each owner. The use of such Third-Party Trademarks is intended to describe the third-party goods or services and does not constitute an affiliation by Intento and its licensors with such company or an endorsement or approval by such company of Intento or its licensors or their respective products or services.

Domains? What are these?

Domain is a corpus from a specific source that may differ from other domains in topic, genre, style, level of formality et cetera*. Basically, a combination of industry sector and content type.

* as defined in "Domain Adaptation and Multi-Domain Adaptation for Neural Machine Translation: A Survey" by Danielle Saunders

Executive Summary

37 Machine Translation Engines evaluated

22 Language pairs

English ⇄

Spanish*
French*
Italian
Portuguese*
German
Dutch

Ukrainian
Korean
Japanese
Chinese*
Arabic

9 Content domains

General
Colloquial
IT

Entertainment
Hospitality
Education

Healthcare
Legal
Financial

* Spanish (LA), French (European), Portuguese (Brazilian), Chinese (Simplified).

23 Machine Translation engines show the best results for some language pairs and domains

- Alibaba E-Commerce
- Alibaba General
- Amazon
- Anthropic
- Baidu
- ChatGPT
- DeepL
- GPT-3
- GPT-4
- Google
- Kawamura
- Microsoft
- Mirai
- ModernMT Static
- Naver
- NiuTrans
- NTT COTOHA
- PROMT
- SYSTRAN
- Tencent
- Ubiqus
- Yandex
- Youdao

Massive language expansion across all MT engines

190K unique language pairs

+65K compared to 2022 — and still growing

654 unique languages

The machine translation market is exploding with MT-capable Large Language Models (LLMs), such as ChatGPT and GPT-4, launched in 2023 by many vendors.

We've evaluated 37 engines overall, among which there are 5 Large Language Models from AI21, Anthropic, and OpenAI.

Several LLMs, such as ChatGPT and GPT-4 by OpenAI, perform on par with the 1st tier commercial engines when used for machine translation.

Translation to English is easier: multiple MT engines perform quite well for all language pairs with English target.

Legal and IT domains, as well as Arabic, Japanese, and Ukrainian languages require a careful choice of MT vendor, as relatively few perform at the top level.

Entertainment and Colloquial domains show relatively low scores, which may indicate the importance for customization there.

About Intento

Intento enables global companies to translate 20x more on the same budget and makes their customer and employee experience instantly multilingual in 650+ languages through Machine Translation and AI.

Intento Enterprise MT Hub augments best-of-breed Custom NMT platforms with Source Quality Improvement, Automatic Post-Editing and Translation Quality Estimation. Combining Machine Translation with Generative AI models (like GPT) enhances traditional translation workflows and achieves accurate, in-context translation.

It integrates with existing software platforms, such as Salesforce, ServiceNow, and Translation Management Systems, to share MT/AI models across the enterprise and improve them based on feedback.

Trusted by the global enterprise

SUBWAY Government of Canada Agilent playrix PROCORE JOOM

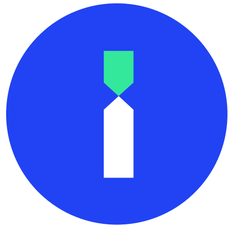
We have been evaluating stock Machine Translation models since May 2017. For customers, we also evaluate customizable NMT models (you can get a glimpse here).

As we show in this report, the Machine Translation landscape is complex and dynamic. Models from five different vendors are required to achieve the best quality in popular language pairs, with a dramatic price difference (as much as 200 times.)

Book a demo

Enterprise Machine Translation Hub

Translate 20x more, with the same budget, across the company – localize content and make customer and employee experience multilingual with real-time machine translation and AI.



Machine Translation Hub for

Book a demo

Localization

Supercharge your TMS to cut backlog and costs using best-of-breed Custom NMT, Source Quality Improvement, and Automatic Post-Editing based on Generative AI. Keep machine translation engines performing at their best with MT Evaluation and MT Maintenance.

crowdin · lokalise · memoq · Trados · XTM International · +more

Customer experience

Scaling means saving, in up to 654 languages. Boost your global multilingual support capacity with agents in any region. Make all your knowledge base articles, community forums, and chatbots work for all markets.

COGNIGY · salesforce · VERINT · zendesk · +more

Employee Experience

Remove language barriers to scale your team globally. Connect the whole company through self-service access to multilingual help materials, chats, and tickets for improved employee-support interactions in any language.

chrome · Confluence · Jira · metricstream · servicenow · Office · +more

About e2f

Established in 2004, e2f helps people and machines understand each other fluently, regardless of language, content, and culture. e2f solutions empower Fortune 50 brands to monitor, objectively assess, and improve communications on a global scale.

e2f delivers world-class translation and training data with its proprietary technology stack for translation, quality review, and AI services. e2f offers a global resource pool of skilled professionals in virtually all countries and languages.

To learn more, contact e2f or visit website.

e2f services

- → MT detection and MT quality evaluation services that enable organizations to monitor suppliers for compliance with brand standards for human and machine translation.
- → Creation of custom Lingosets™, or augmented multilingual datasets that represent real human conversational flow. Lingosets serve as benchmarks for conversational AI deployments.
- → Golden datasets and training datasets that enable leading MT providers to evaluate and fine-tune engine performance.

Overview

1. MT Engines
2. Datasets
3. Evaluation Methodology
4. Evaluation Results
5. Miscellaneous
6. Takeaways

37 Machine Translation Engines

22 Language Pairs

9 Content Domains

1. MT Engines

1.1 Machine Translation Landscape

Generic stock models

- AISA
- Alibaba
- Amazon
- AppTek
- Baidu
- DeepL
- eBay
- Elia
- Fujitsu
- Globalese
- Google
- GTCom
- IBM
- iFlyTek
- Kakao
- Kawamura powered by NICT
- Kingsoft
- Lesan
- Lindat
- LingvaNex
- Microsoft
- Mirai
- ModernMT
- Naver
- NiuTrans
- NTT COTOHA
- Omniscien
- Oracle
- PangeaMT
- Phrase
- Process9
- Prompsit
- PROMT
- Reverso
- Rozetta
- RWS
- Sogou
- SYSTRAN
- Tencent
- TREBE
- Tilde
- Ubiqus
- Unbabel
- Yandex
- YarakuZen
- Youdao

Vertical Stock Models

- Alibaba
- Baidu
- CloudTranslation
- Microsoft
- NiuTrans
- Omniscien
- PROMT
- RoyalFlush
- SAP
- SYSTRAN
- Ubiqus
- XL8

Custom terminology support

- Amazon
- Baidu
- DeepL
- Google
- IBM
- Microsoft
- Rozetta
- RWS
- SYSTRAN
- Ubiqus
- Yandex

Static domain adaptation

- Alibaba
- AppTek
- Baidu
- CloudTranslation
- Globalese
- Google
- IBM
- KantanAI
- Microsoft
- Omniscien
- PangeaMT
- Prompsit
- PROMT
- Rozetta
- RWS
- SYSTRAN
- Tilde
- Ubiqus
- Yandex

Dynamic domain adaptation

- Amazon
- ModernMT

Large Language Models

- AI21
- Anthropic
- BigScience
- Cohere
- EleutherAI
- Google
- LAION
- Microsoft
- MosaicML
- OpenAI
- Stability

Standalone commercial products with an API. All product names, trademarks and registered trademarks are property of their respective owners. All company, product and service names used in this material are for identification purposes only. Use of these names, trademarks and brands does not imply endorsement.

1.1 Machine Translation Landscape

Generic Stock Models

Pre-trained models based on data from multiple sources. These models are not pre-adjusted to one particular industry or specialization, such as Legal or Medical translations.

Custom Terminology Support

Allows users to customize the MT models by applying their own glossaries. Depending on the implementation, terminology can be used while training custom models or for adjusting machine translation results.

Dynamic Domain Adaptation

The model can be incrementally updated on the fly. The adaptation can be done with as few as a single datapoint and happens in real-time. Typically, there's no snapshot of the baseline model created, making the model performance affected when the baseline model is updated by an MT provider.

Vertical Stock Models

Pre-trained models, pre-adjusted to one particular industry or specialization, such as Legal or Medical translations.

Static Domain Adaptation

The baseline MT model can be adjusted using batch training. The training requires a significant amount of data (thousands of parallel segments) and takes time (from hours to days). Once the model is trained, a snapshot of a model is created and does not change after the next batch re-training.

Large Language Models

Large Language Models (LLMs) are trained on massive amounts of data to generate text, follow instructions, and answer questions. These models can be used for various tasks such as content creation, sentiment analysis, text summarization, or translation.

1.2 Evaluated Machine Translation Engines

Customization options

- None
- TM
- Glossary
- Both

AI21 Generative text AI ●	Alibaba eCommerce MT ○	Alibaba Cloud General ○	Amazon Translate ●	Anthropic Next-generation AI assistant ●
Apptek Neural Machine Translation ○	Baidu Translate API ◒	Cloud Translation Translation API ○	NTT COTOHA Translator ○	DeepL API ◒
Elia Elhuyarren itzult-zaile automatikoa ○	Globalese Machine Translation ◒	Google Cloud Advanced Translation ●	IBM Watson eCommerce MT ●	Kawamura by NICT Translation Engine ○
Microsoft Language Translator ●	ModernMT Static ◒	Mirai Translator ○	Naver Papago NMT Commercial ○	NiuTrans Translation Cloud Platform ○
NVIDIA NeMo framework ◒	OpenAI GPT LLM x3 ●	Oracle Machine Translation ○	Pangeanic Machine Translation API ○	PROMT Cloud API ○
RoyalFlush Finance Translation ○	SAP Machine Translation ○	SYSTRAN PNMT ●	TartuNLP Neurotõlge MT ○	Tencent Cloud TMT API ○
Tilde Machine Translation API ○	TREBE Machine Translation API ○	Ubiqus Translation API ●	Yandex Translate API ●	Youdao Cloud Translation API ○

Large Language Models can be customized with TMs through fine-tuning, and terminology via prompt engineering

2. Datasets

2.1 Preparation

The source data collection and initial cleaning were done by Intento.

Open-Source English Texts

Carefully selected from open-source data

→ Found several resources for each domain and selected the ones with suitable license agreements

→ Extracted high-quality segments

Data samples for various domains are used according to their licence agreements: Financial data, Hospitality data 1, Hospitality data 2, Legal data, Entertainment data, IT data, Colloquial data

Filtering to Ensure High-Quality Source

Collected data for 9 domains using open-source resources

→ Removed duplicates, tags, and broken symbols

→ Removed segments under 4 words

→ Removed segments that were truncated (except for the Colloquial sector) and segments that were longer than one sentence

→ Manually checked each segment in every domain to avoid segments with an ambigous meaning or incorrect tone of voice

2.1 Preparation

The dataset translations and quality assurance were done by e2f.

Translation by Native Speaking Experts

→ Selected native translators with expert-level qualifications and positive feedback in each language and domain.

→ For reviews, selected native language experts in editing and proofreading across multiple domains, and positive customer feedback.

→ Proofread strings supplied by Intento for compliance with proper English grammar, spelling, and punctuation and supplied files to translators via e2f's Translation, Editing, and Proofreading (TEP) platform.

Quality Assurance

Provided via e2f's TEP portal

→ Human translations were compared with ones generated by the leading machine translation engines using e2f's MT Detection tool, and determined the probability that they contained machine-translated and/or post-edited content (MTPE).

→ Strings whose MTPE probability exceeded e2f's threshold triggered expert review and was followed by re-translations, which were automatically reassessed. The resulting golden dataset does not bear traces of MTPE.

→ Quality assurance reports were run on capitalization, punctuation, spelling, numbers, spaces, and typos. Reviewers implemented necessary changes and proofread the dataset prior to final delivery.

2.1 Preparation

Additional cleaning was done after MT Report 2022

Additional cleaning

→ Deleted segments with issues in the reference reported by our linguists

→ Deleted segments with lowest scores in each domain+language pair dataset using QE metrics to evaluate reference translation quality

→ Deleted full forms of acronyms in the reference sentences that were not present in the source sentences

Reversing the direction to English

→ Reversed data to Language X > English

→ Deleted segments with human translations being “too localized” since they could lead to scores corruption

→ Deleted segments with scores lower for reference than for MT in each domain+language pair dataset using QE-metrics and back-translations

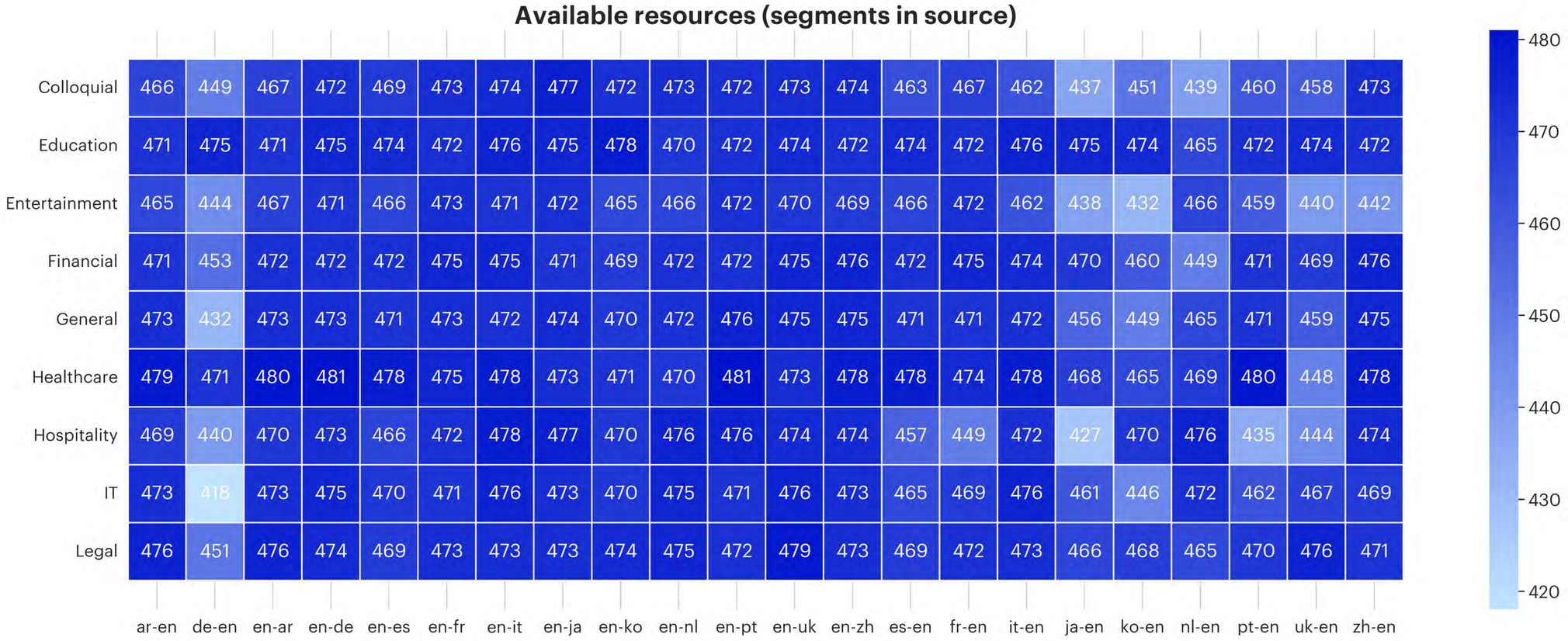
2.2 Content Domains and Language Pairs

9

content domains
per language pair

22

language pairs
per domain



2.3 Content Samples by Domain

General

“Walmart is also the largest grocery retailer in the United States.”

Finance

“Both operating profit and net sales for the three-month period increased, respectively from €16m and €139m, as compared to the corresponding quarter in 2006.”

Hospitality

“Very reasonably priced and the food is excellent, I had pasta which was delicious, and my friend had the Italian meats & cheeses.”

Healthcare

“Leishmaniosis caused by Leishmania infantum is a parasitic disease of people and animals transmitted by sand fly vectors.”

Legal

“Landlord and Tenant acknowledge and agree that the terms of this Amendment and the Existing Lease are confidential and constitute proprietary information of Landlord and Tenant.”

Entertainment

“Further, they are aided by a magnificent cast of co-stars, most notably their secretary, played by Isabel Tuengerthal, who is a rare gem with great comic potential.”

Education

“Find what straight lines are represented by the following equation and determine the angles between them.”

IT

“The interface is in Python, a dynamic programming language, which is very appropriate for fast development, but the algorithms are implemented in C++ and are tuned for speed.”

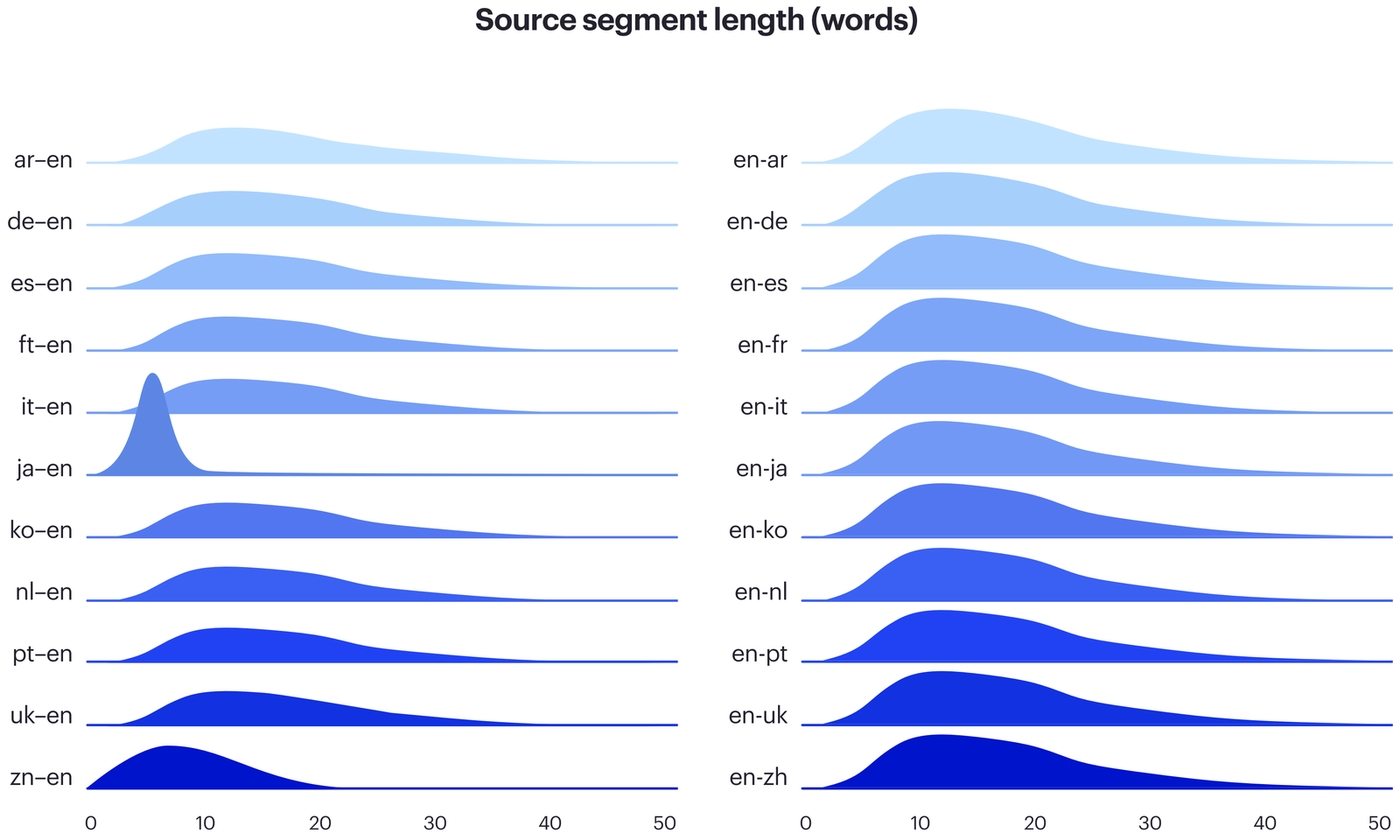
Colloquial

“and, in fact, there are two huge lenses that frame the figure on either side”

2.4 Sentence Length

→ 22 language pairs were translated in total: 11 pairs from English into other languages, and 11 pairs into English.

→ Sentences that were too short (< 4 words) were excluded from the dataset.



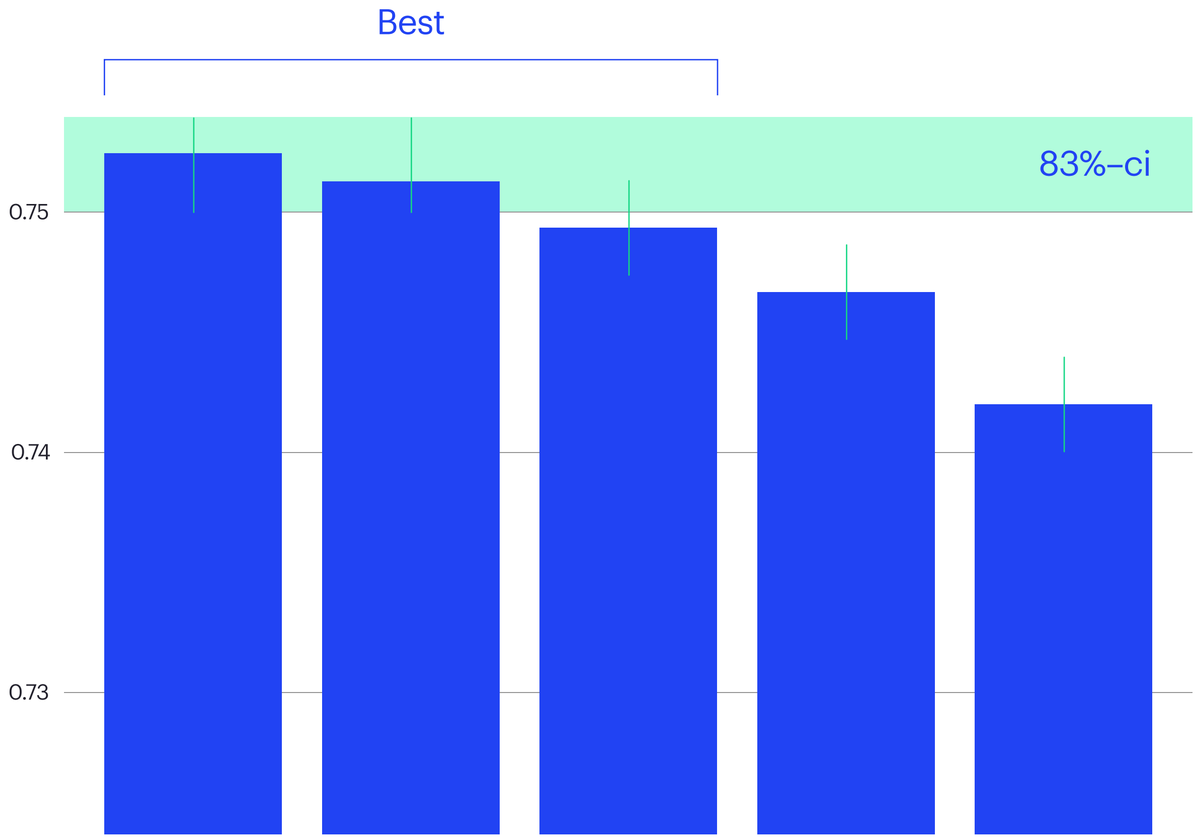
3. Evaluation Methodology

3.1 Evaluation Approach

1. Rank MT engines based on a score showing distance from a reference human translation.

2. Identify a group of top-runners (BEST) within a confidence interval of the leader.

Using segment-level scores averaged across the corpus and an 83% confidence interval [1,2]



1. Harvey Goldstein; Michael J. R. Healy. The Graphical Presentation of a Collection of Means, Journal of the Royal Statistical Society, Vol. 158, No. 1. (1995), p. 175-177.

2. Payton ME, Greenstone MH, Schenker N. Overlapping confidence intervals or standard error intervals: what do they mean in terms of statistical significance?. J Insect Sci. 2003;3:34. doi:10.1093/jis/3.1.34.

3.2 What Scores to Use

hLEPOR

Syntactic similarity

Compares similarity of token-based n-grams. Penalizes both omissions and additions. Penalizes paraphrases / synonyms. Penalizes translations of different length.

paper + code

TER

Syntactic similarity

Measures the number of edits (insertions, deletions, shifts, and substitutions) required to transform a machine translation into the reference translation. Penalizes paraphrases/ synonyms. Penalizes translations of different length.

paper + code

SacreBLEU

Syntactic similarity

Compares token-based similarity of the MT output with the reference segment and averages it over the whole corpus. Penalizes omissions and additions. Penalizes paraphrases / synonyms. Penalizes translations of different length.

paper + code

BERTScore

Semantic similarity

Analyzes cosine distances between BERT representations of machine translation and human reference (semantic similarity). Does not penalize paraphrases / synonyms. May be unreliable for terminology in domains and languages underrepresented in BERT model.

paper + code

☆ COMET

Semantic similarity

Predicts machine translation quality using information from both the source input and the reference translation. Achieves state-of-the-art levels of correlation with human judgement. May penalize paraphrases/synonyms.

paper + code

3.3 MT Using Large Language Models

- → On November 28, 2022, OpenAI released their newest GPT-3 model, davinci-003. In January 2023, we have evaluated GPT-3 for machine translation scenario.
- → In March 2023, we have evaluated ChatGPT (gpt-3.5-turbo) and GPT-4 for machine translation, as well.
- → Anthropic and AI21 are new to Intento and are evaluated for the first time in this Report.
- → We have performed several experiments with prompting the Large Language Models to determine the suitable prompt for the Report, described in the articles referenced above.
- → For this evaluation, we used a system message mentioning a specific domain, for example, "*You are a professional translator working with software documentation.*", and a prompt that asks the model to perform translation, i.e. "*Translate this from <source language> to <target language>: < source segment>*", which consistently work for both completion and chat-style models.

4. Evaluation Results

4.1 Best MT Engines per Language Pair (COMET)

- → 5 MT engines are among the statistically significant leaders for 22 language pairs.
- → Amazon, DeepL, and Google cover the best options for all languages when domains are ignored.
- → Higher linguistic quality can be achieved using engine customization and glossary support.
- → Absolute values are not shown to avoid confusion, as the scores are not comparable across language pairs.
- → The domain and content type mix is differenent for every language pair (see the next slide) and largely influences this leaderboard.

* Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

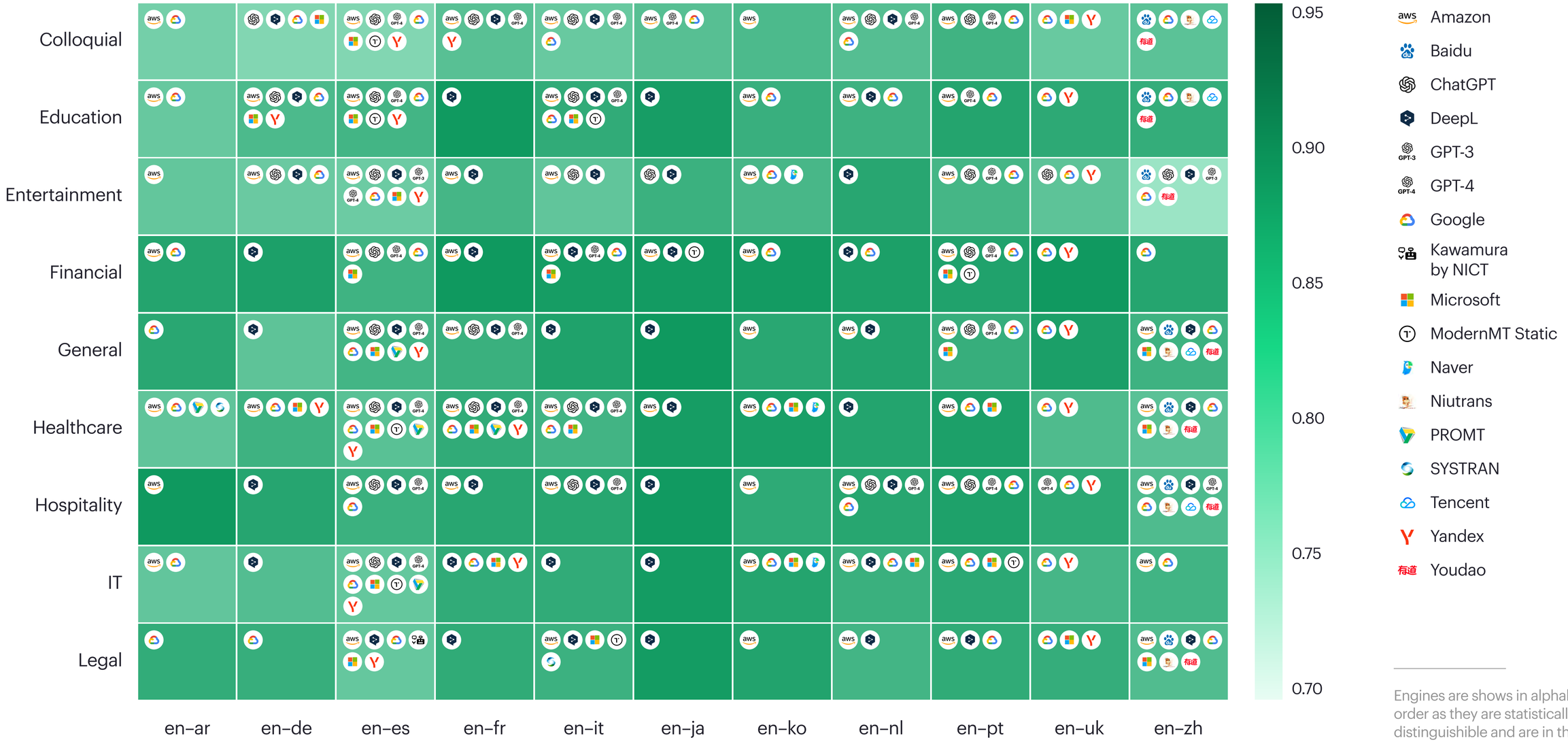
Best MT engines by normalized COMET score*

ar–en		en–uk	
de–en		en–zh	
en–ar		es–en	
en–de		fr–en	
en–es		it–en	
en–fr		ja–en	
en–it		ko–en	
en–ja		nl–en	
en–ko		pt–en	
en–nl		uk–en	
en–pt		zh–en	

4.2 Best MT Engines per Domain. English Source

- In the next slide, we show the best MT engines by normalized COMET score. Each square shows the best providers for a particular language pair in a specific domain. The color of the square shows the achievable MT quality for this domain compared to other domains in this language pair.
- For example, we see that the best engine for the English-German pair in the Financial and IT domains is DeepL. Its score for the Financial domain is higher, so we expect less post-editing than in the IT domain.
- For each language pair, the score values were normalized to the [0,1] range, hence it's not comparable between different language pairs.
- MT vendors in one bucket provide the best quality for this language pair and domain, with no statistically significant difference between them. They are presented in alphabetical order.

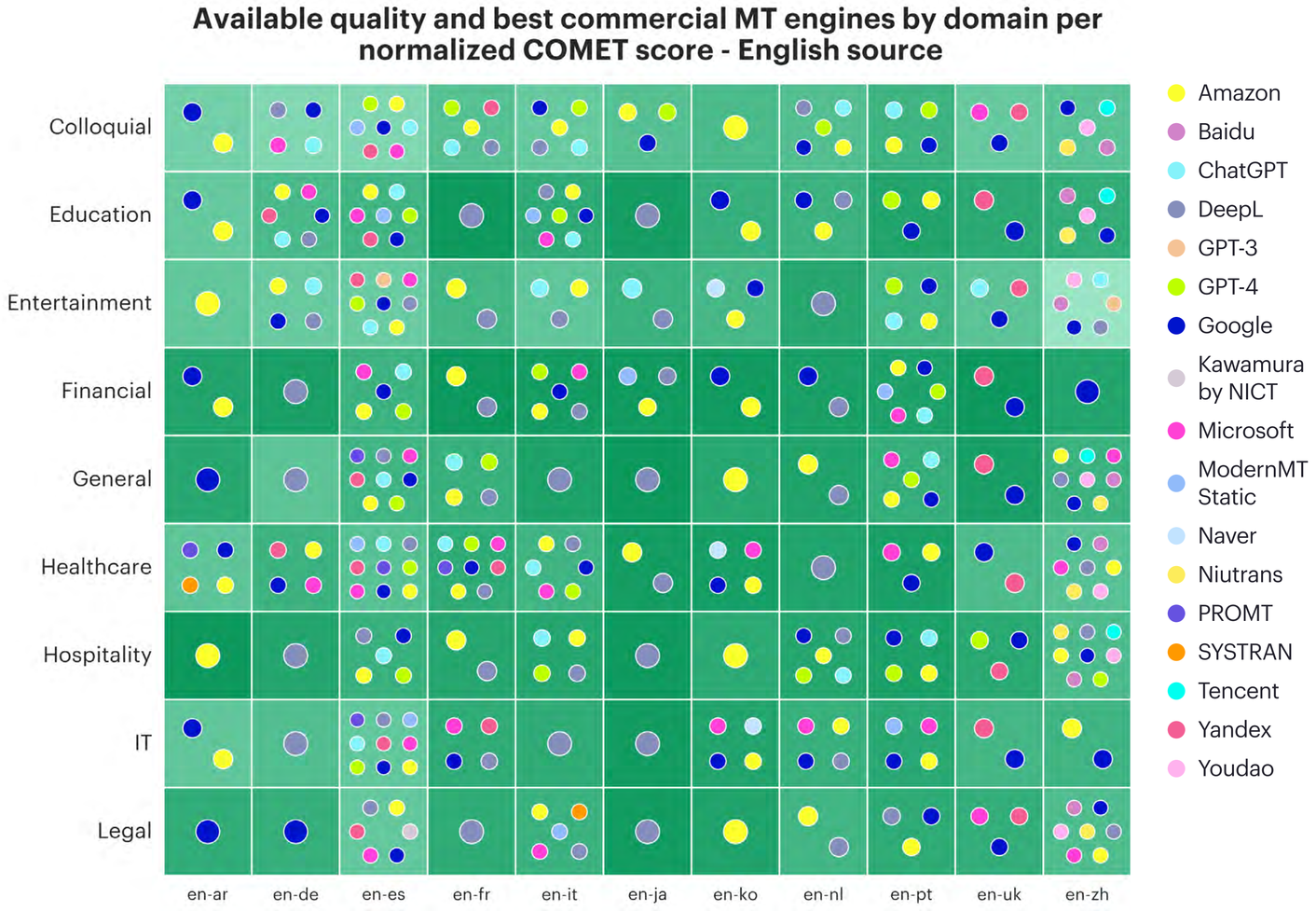
Available quality and best commercial MT engines by domain per normalized COMET score
English source



Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

4.2 Best MT Engines per Domain. English Source

- → 17 MT engines are among the statistically significant leaders for 9 domains and 11 pairs with English source.
- → Many engines perform best with English to Spanish, Portuguese, and Chinese.
- → Legal and IT domains, as well as Arabic and Japanese languages, require a careful choice of MT vendor, as relatively few perform at the top level.
- → Despite having several comparable engines per language pair, Entertainment and Colloquial domains show relatively low scores, which may indicate the importance of customization.
- → Several LLMs are performing on par with the commercial systems: GPT-3, ChatGPT (gpt-3.5-turbo), and GPT-4.



4.3 Possible Minimal Coverage. English Source

5 MT engines provide minimal coverage* for all pairs and industries, 2–3 per domain.

Entertainment
Amazon, ChatGPT, DeepL

Hospitality
Amazon, DeepL, GPT-4

Legal
Amazon, DeepL, Google

General
Amazon, DeepL, Google

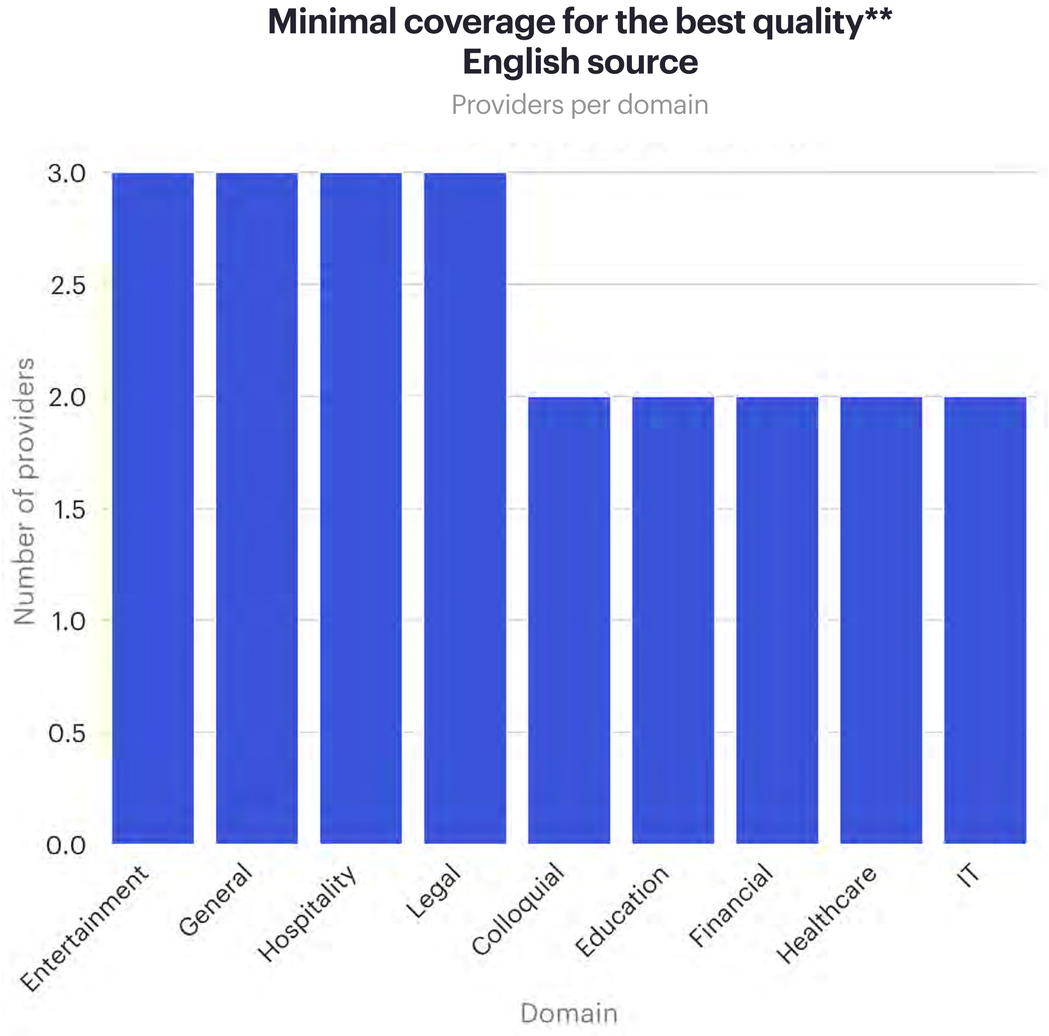
Colloquial
Amazon, Google

Education
DeepL, Google

Healthcare
DeepL, Google

IT
DeepL, Google

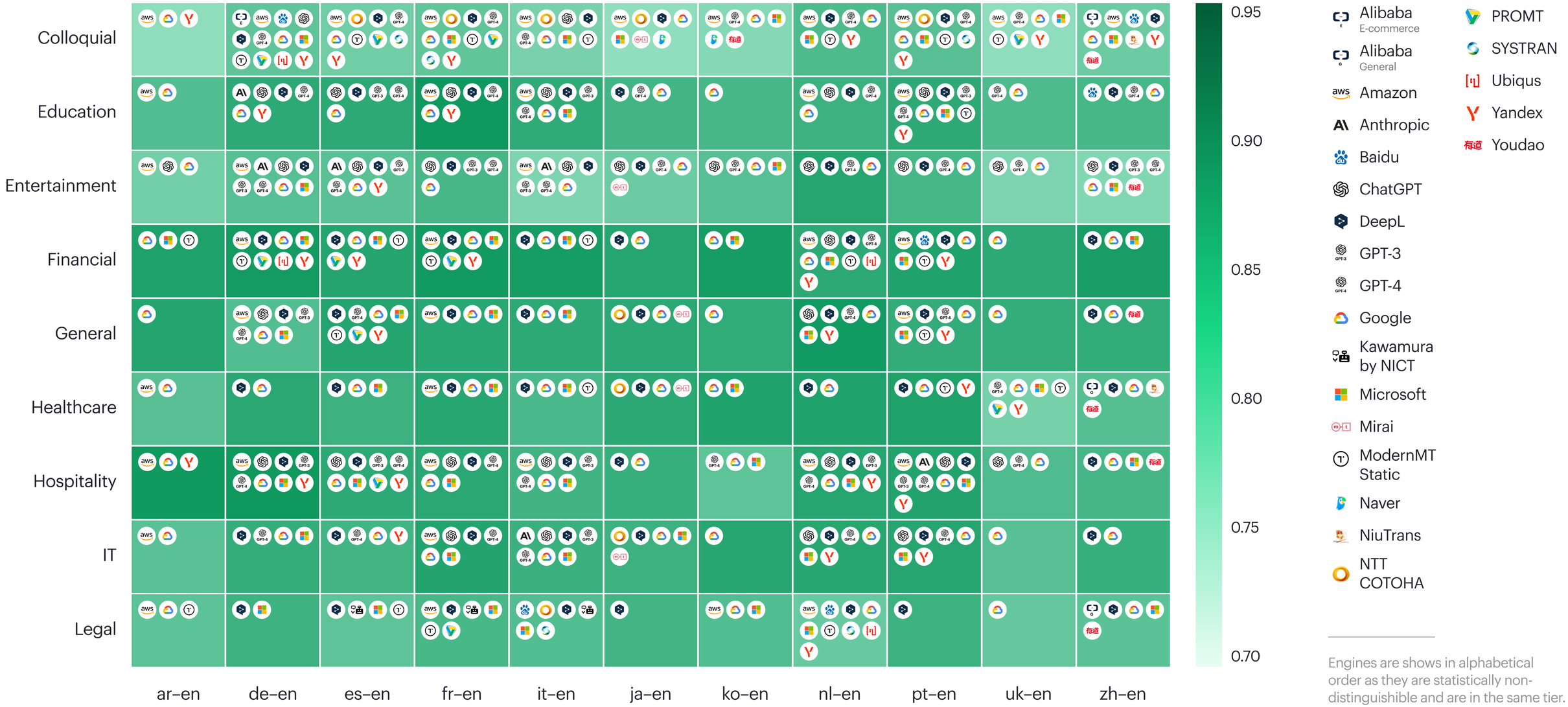
Financial
DeepL, Google



* For every domain, we provide the minimum number of providers needed to translate all language pairs in this specific domain.

** Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

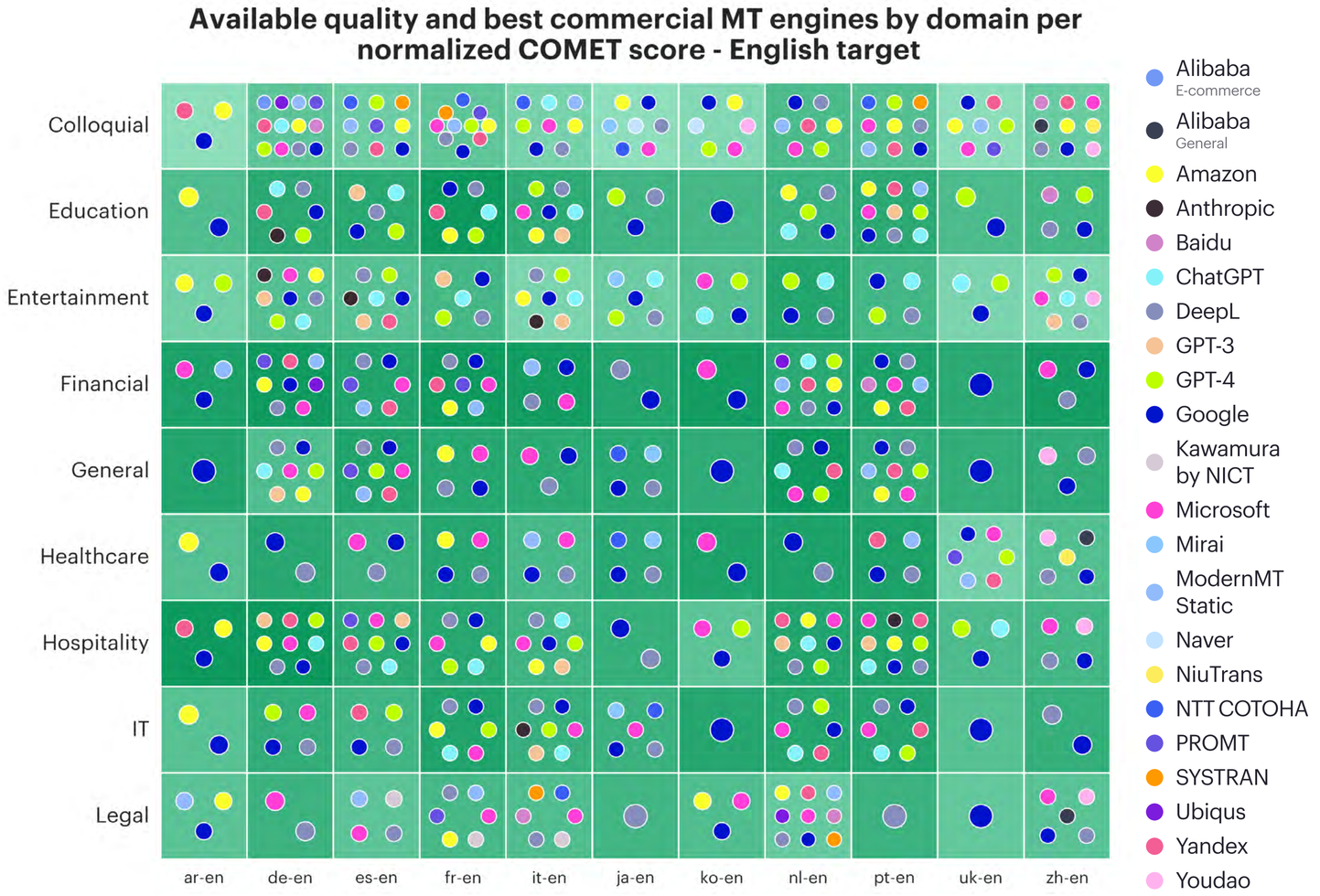
Available quality and best commercial MT engines by domain per normalized COMET score
English target



Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

4.4 Best MT Engines per Domain. English Target

- → 22 MT engines are among the statistically significant leaders for 9 domains and 11 pairs with English target.
- → Most pairs with English target have a large number of highly performing engines.
- → Ukrainian to English require a careful choice of MT vendor, as relatively few perform at the top level.
- → Despite having several comparable engines per language pair, Entertainment and Colloquial domains show relatively low scores, which may indicate the importance of customization.
- → Several LLMs are performing on par with the commercial systems: Anthropic, GPT-3, ChatGPT (gpt-3.5-turbo), and GPT-4.



4.5 Possible Minimal Coverage. English Target

4 MT engines provide minimal coverage* for all pairs and industries, 2–3 per domain.

Legal
DeepL, Google

Entertainment
GPT-4

Colloquial
Amazon

Healthcare
Google

Financial
Google

General
Google

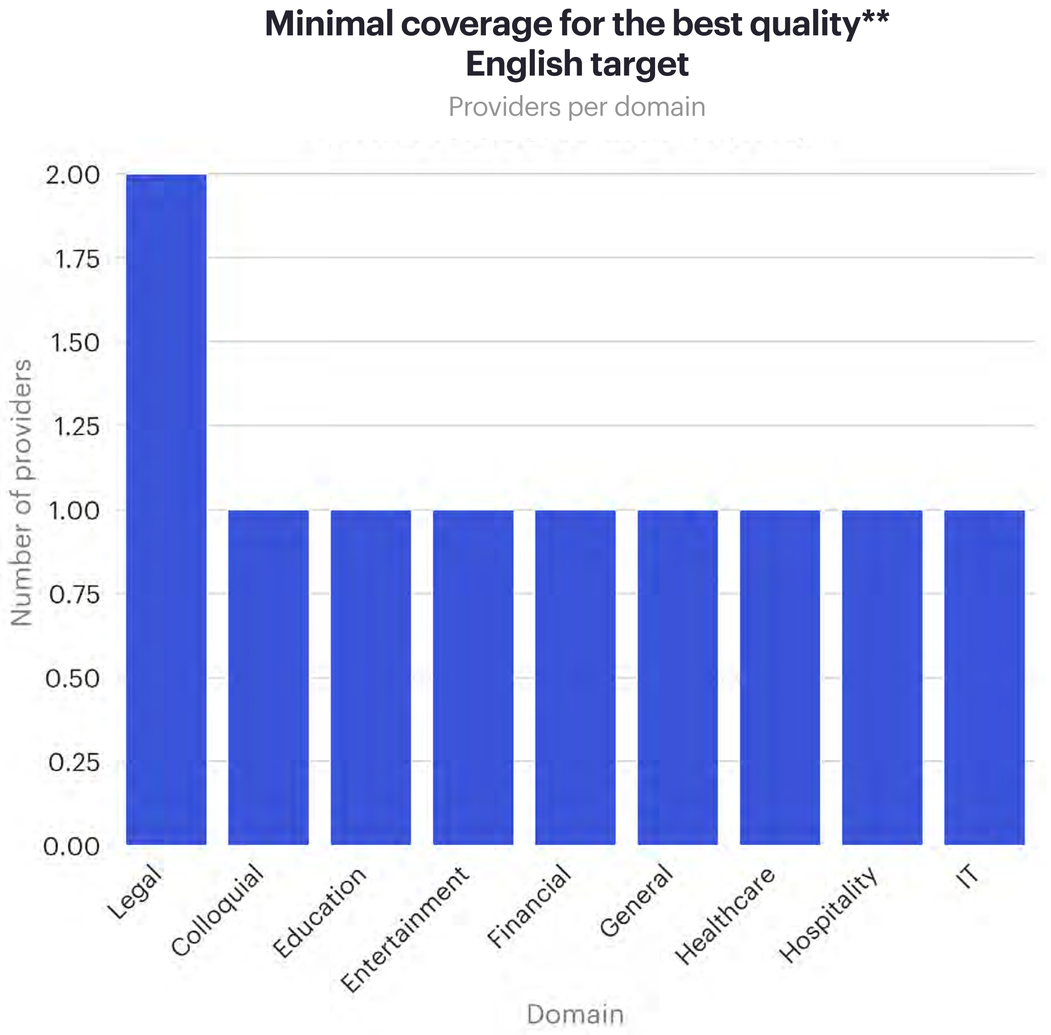
Hospitality
Google

Education
Google

IT
Google

* For every domain, we provide the minimum number of providers needed to translate all language pairs in this specific domain.

** Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

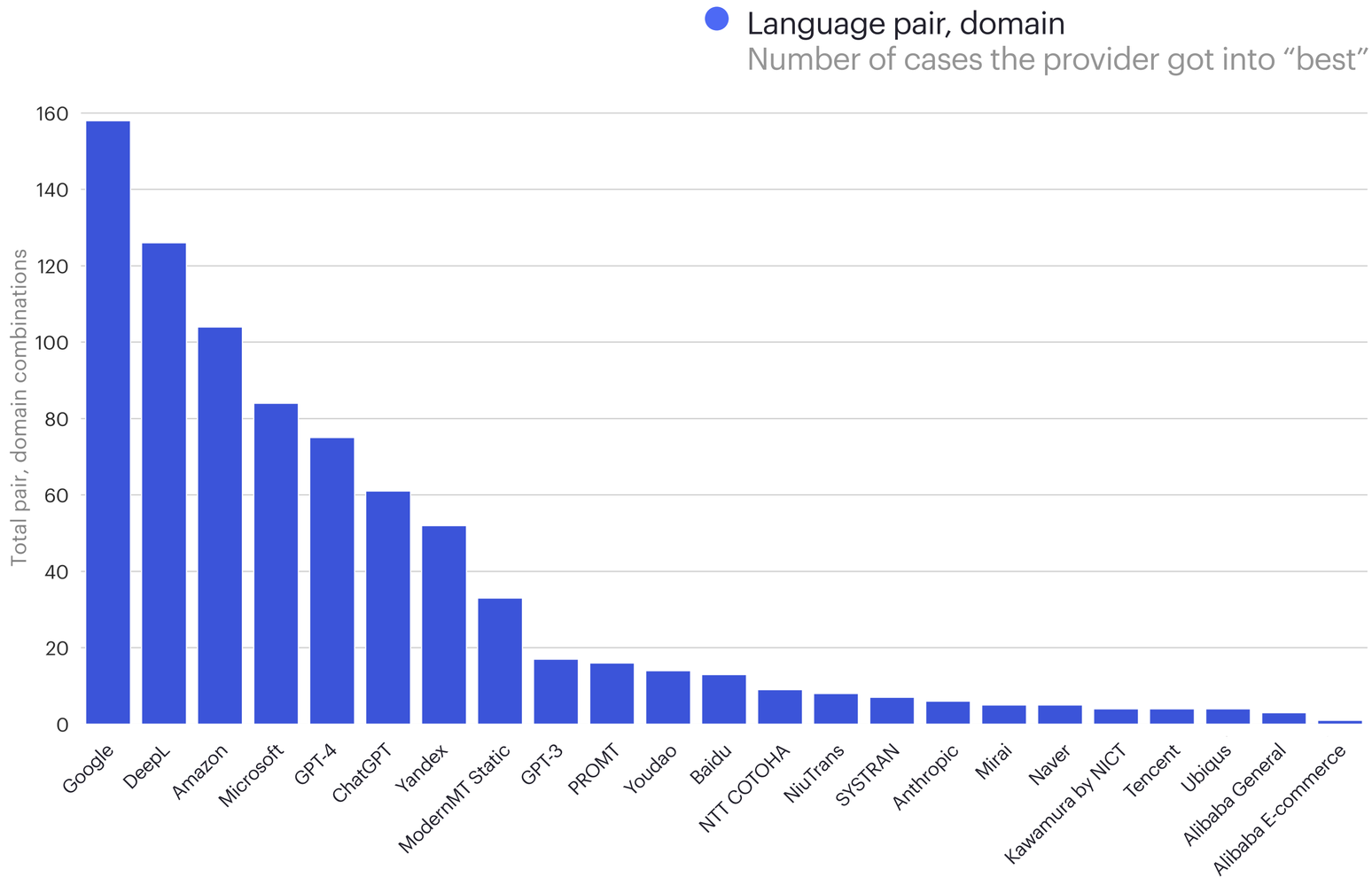


4.6 Top Performing MT Providers (COMET)

22 language pairs, 9 domains

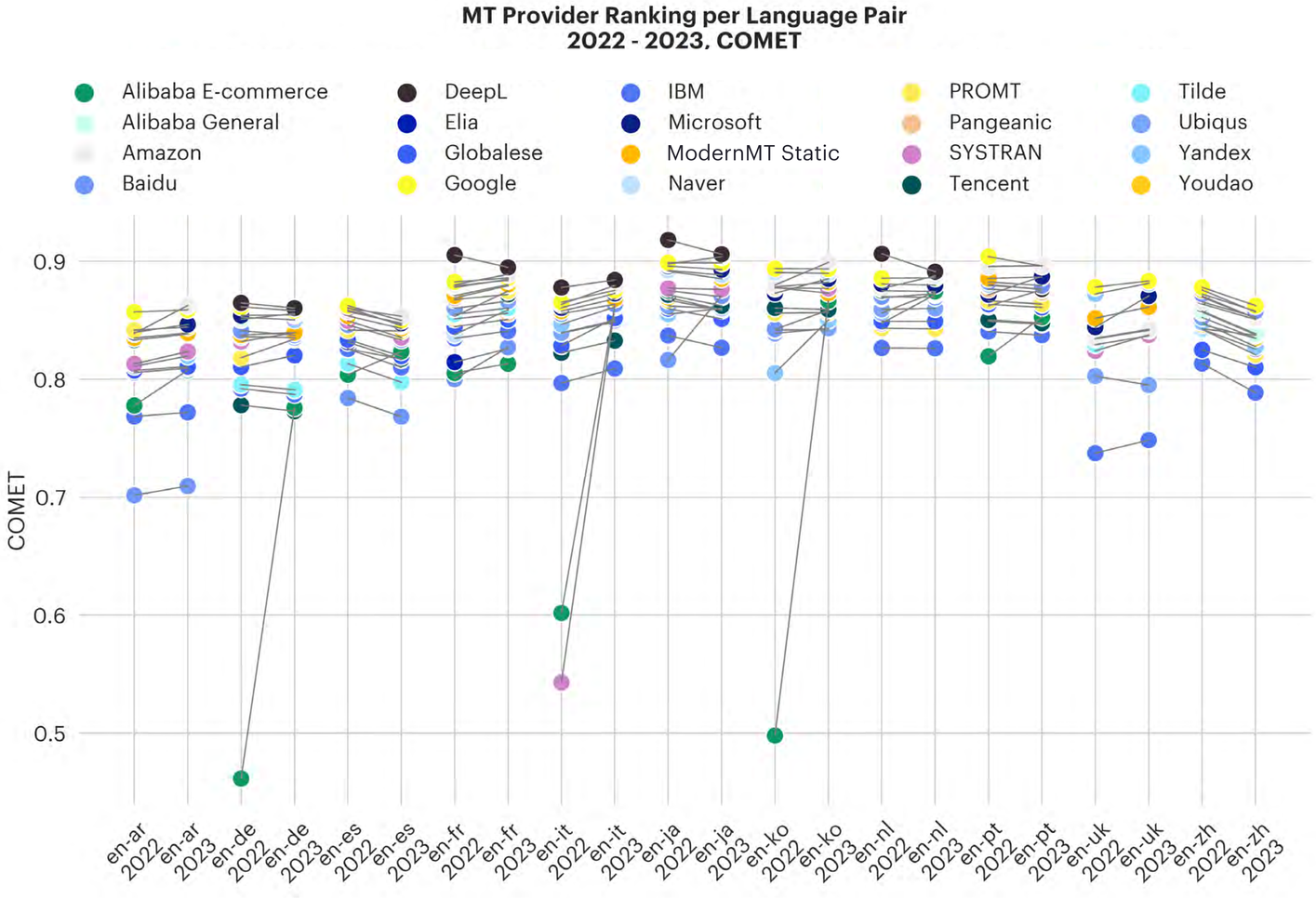
Some providers were tested only in their specific domains and language pairs:

- → HiThink RoyalFlush specializes in en-zh translation in the Finance domain
- → TREBE specializes in Iberian languages, and was used for en-es and es-en translation
- → SAP is intended to be used with “SAP-specific” tasks and is not to be compared with the genetic stock models
- → Out of different ModernMT models, the static model was evaluated



4.7 Historical results (COMET)

- → Alibaba E-commerce has improved significantly in pairs en-de, en-it, and en-ko.
- → DeepL has lower COMET scores than in the previous year. In en-ja, we observe some moderate to severe omissions and mistranslations. Overall, we see a minor decrease in scores in several pairs except en-it.
- → We observe lots of improvements for en-it, some of them are quite significant (SYSTRAN, Alibaba E-commerce, and ModernMT).
- → A large number of providers have lower COMET scores for en-zh and en-es compared to the previous year.



5. Miscellaneous

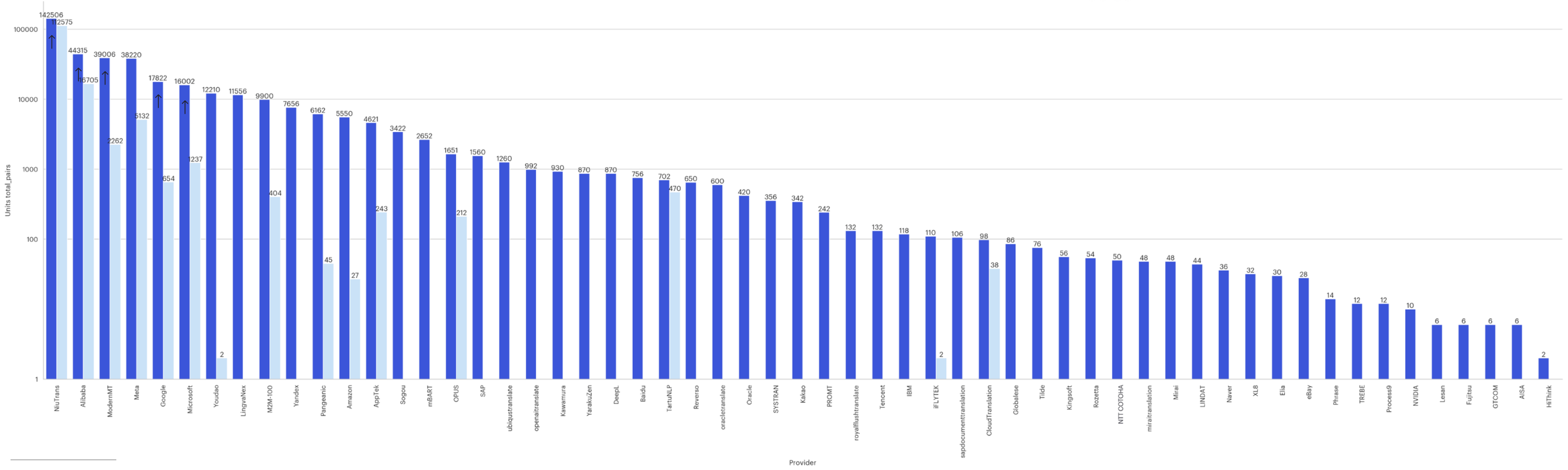
5.1 190,085 Language Pairs Across All MT Engines*

- total language pairs
- unique language pairs
- ↑ language pair growth

From 125,075 in July'22 to 190,085 in May'23

Significant growth for NiuTrans, Google, ModernMT, and Alibaba

Added new niche MT providers with few languages



* Where possible, we have checked via API if all language pairs advertised by the documentation are supported and removed the pairs we were unable to locate in the API.

** As advertised (not validated via API).

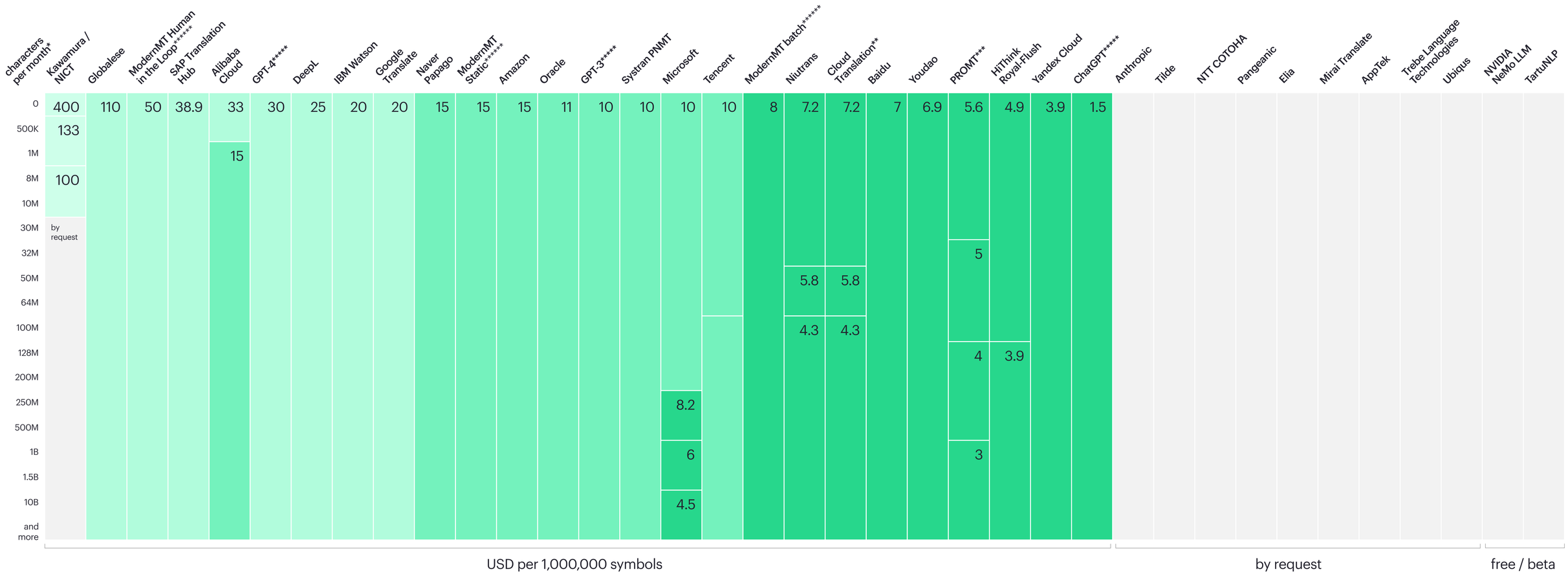
*** Due to LLMs multilinguality and the nature of the data they were trained on, there is no definitive list of languages or language combinations they support. For now, we do not include LLMs in this slide.

5.2 Changes in Providers' Features

- → ModernMT added more than a 100 new languages to their platform, making it an overall count of 200 languages, among which there are 30 languages that are available commercially the first time.
- → DeepL added several more languages to its glossaries feature, and is now providing support for glossaries in any language combination from 8 languages overall.
- → NiuTrans added 78 language pairs to its API, among which 74 are considered low-resource and were added to Intento for the first time.
- → IBM is announcing the deprecation of the IBM Watson® Language Translator service for IBM Cloud® in all regions. As of 10 December 2024, the service will be withdrawn entirely and will no longer be available to any customers
- → In March 2023, OpenAI has announced their latest AI contribution, GPT-4. GPT-4 can work on various tasks with greater accuracy, because of its vast general knowledge and problem solving abilities.
- → In May 2023, Google has introduced their latest AI solution, PaLM 2. PaLM 2 is described as a state-of-the-art large language model that has improved multilingual, reasoning and coding capabilities.

5.3 Public Pricing. Stock Models

USD per 1M characters****



* Prices provided herein are based on the publicly listed prices at the time of the analysis. Actual prices may vary depending on a variety of factors, including your geographical location and any customary discounts. It is always recommended to contact the vendor directly for the most accurate and up-to-date pricing information.

** Volume estimation based on 4.79 characters per word.
*** +20% for some language pairs.
**** Freemium volumes are not shown.

***** Prices are converted with an estimation of 2.83 characters per token.
****** From three ModernMT offering, ModernMT Static was evaluated in this report.

5.3 Public Pricing. Domain Adaptation

		Amazon	Google v3	Globalese	GPT-3	IBM v3	Microsoft v3	ModernMT Human-in-the-loop	SYSTRAN PNMT
Training		free	$45/Hour, max $300	$54.90	$11 per million source + target chars of training data for 1 training epoch**	by request	$10 per million source + target chars of training data (max. $300/ training)	free	by request
Hosting/month		200 GB free parallel data storage for every account. Customers pay $0.023 per GB per month for excess data stored.	free	$5.50	free	by request	$10	free	by request
Translation*	1M	$60	$80	$110	$60*	$80	$40	$50	by request
	>250M		$60						
	>2.5B		$40						
	>4B		$30						

* Prices provided herein are based on the publicly listed prices at the time of the analysis. Actual prices may vary depending on a variety of factors, including your geographical location and any customary discounts. It is always recommended to contact the vendor directly for the most accurate and up-to-date pricing information.

** Prices are converted with an estimation of 2.83 characters per token.

5.4 Independent Cloud MT Vendors with Stock Models

Commercial **48**

AISA, Alibaba, Amazon, Apptek, Baidu, CloudTranslation, DeepL, Elia, Fujitsu, Globalese, Google, GTCom, IBM, iFlyTek, RoyalFlush, Lesan, Lindat, Lingvanex, Kantan, Kawamura / NICT, Kingsoft, Masakhane, Microsoft, Mirai, ModernMT, Naver, Niutrans, NTT, Omniscien, Pangeanic, Prompsit, PROMT, Process9, Reverso, Rozetta, RWS, SAP, Sogou, Systran, Tencent, Tilde, Ubiqus, Unbabel, TREBE, XL8, Yandex, YarakuZen, Youdao

Preview / Limited **5**

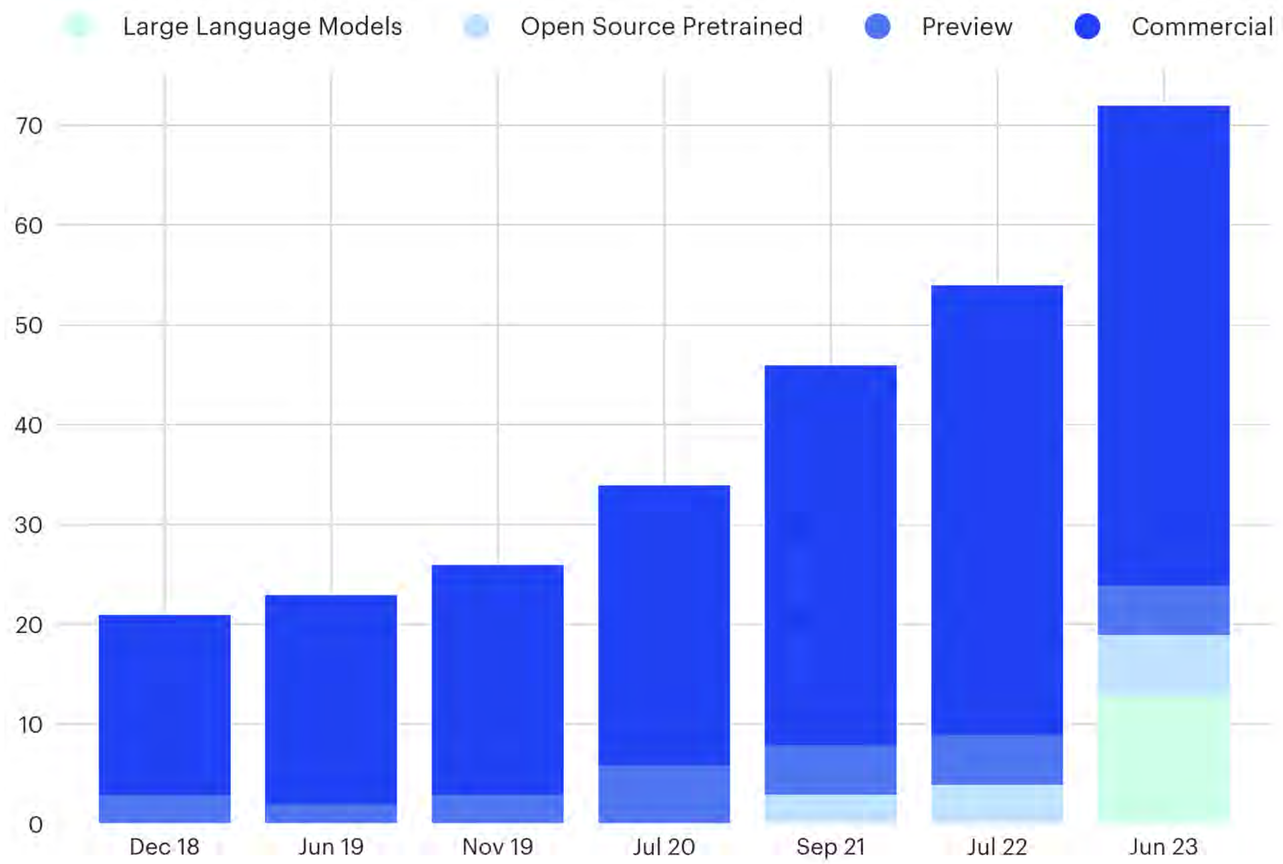
eBay, Kakao, QCRI, Tarjama, Birch.AI

Open Source Pretrained **6**

TartuNLP, NeMo by NVIDIA, NLLB by Meta AI, M2M-100, mBART, OPUS

Large Language Models **13**

AI21, Anthropic, BAIR, BigScience, Cerebras, Cohere, DataBricks, EleutherAI, Google, Meta AI, MosaicML, OpenAI, Stability



The new engines are highlighted in blue.

5.5 Open Source Pre-Trained MT Engines

Neurotõlge by TartuNLP

code

Neurotõlge is a multidirectional machine translation engine developed by the NLP lab at the University of Tartu. Among the high-resource languages, such as English or German, the engine supports several low-resource languages from the Finno-Ugric language family, such as Võro, Livonian, Moksha, etc.

NeMo by NVIDIA

NVIDIA NeMo is a conversational AI toolkit built for researchers working on various AI tasks, such as automatic speech recognition, large language models, and natural language processing.

OOS models evaluated in the 2022 MT Report

In the last year's "State of the Machine Translation", we evaluated NLLB by Meta AI.

We have decided to omit NLLB models in this year's report as they have only shown results in the 2nd tier of commercial systems.

OOS models evaluated in the 2021 MT Report

In the 2021's "State of the Machine Translation", we evaluated three open-source models: OPUS MT, M2M-100, and mBART50.

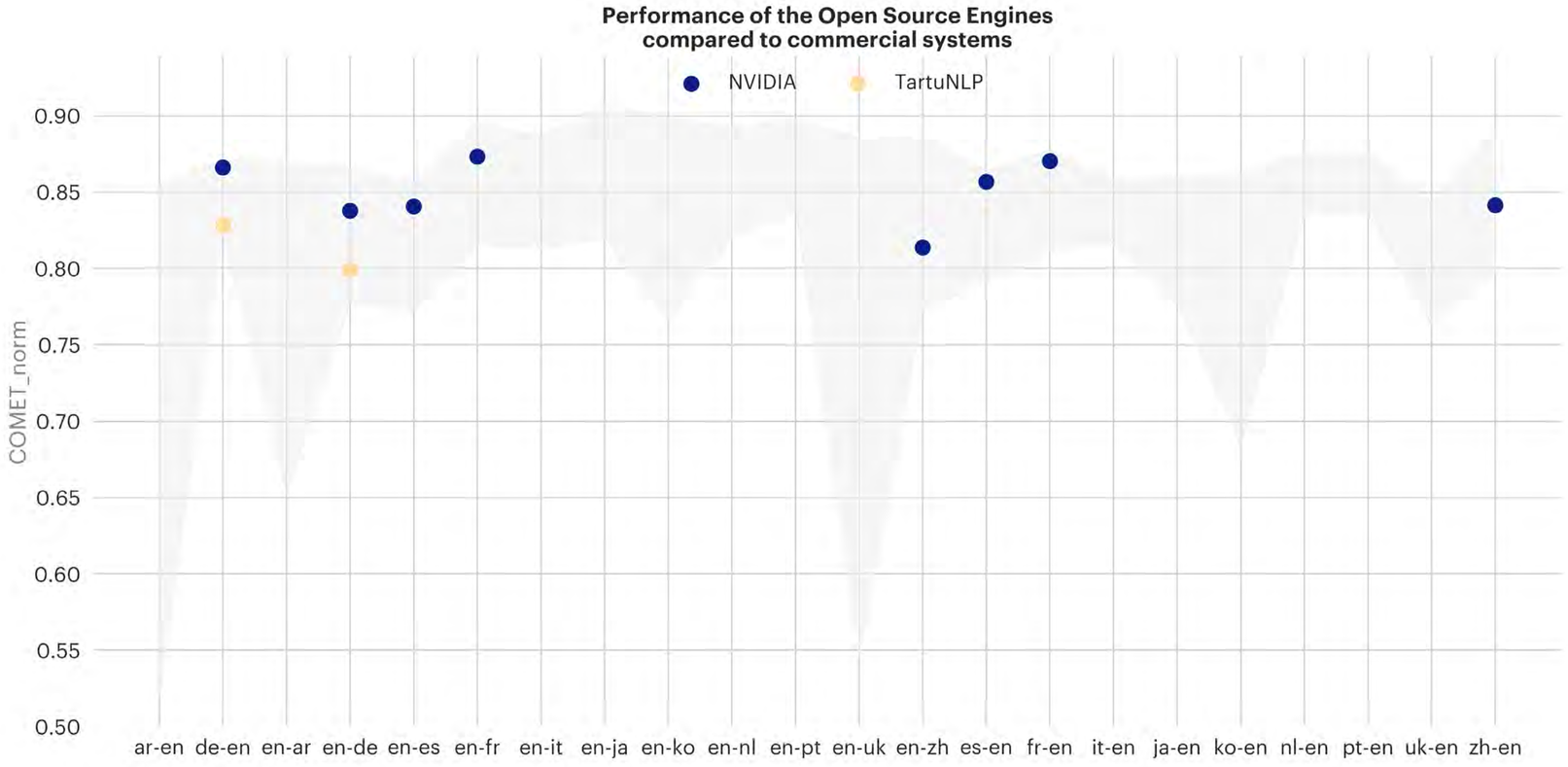
We have decided to omit them in this year's report as they have only shown results in the 2nd tier of commercial systems.

5.6 Open Source MT Performance (COMET)

→ NVIDIA's NeMo machine translation model performs on par with commercial systems when it comes to pairs with English target.

→ However, it only scores in the second tier in pairs with English source.

→ TartuNLP performs in the second tier compared to commercial systems in the en<>de pairs represented in the MT Report.



5.7 Large Language Models

GPT models by OpenAI

GPT models are neural network-based large language models. They are capable of working with a large number of natural language processing tasks, among which there is machine translation. GPT-3 is a text completion model, ChatGPT (gpt-3.5-turbo) is fine-tuned for chat conversation from a model in the GPT-3.5 series, and GPT-4 is the next generation of models after GPT-3/3.5 with even higher capabilities.

Claude by Anthropic

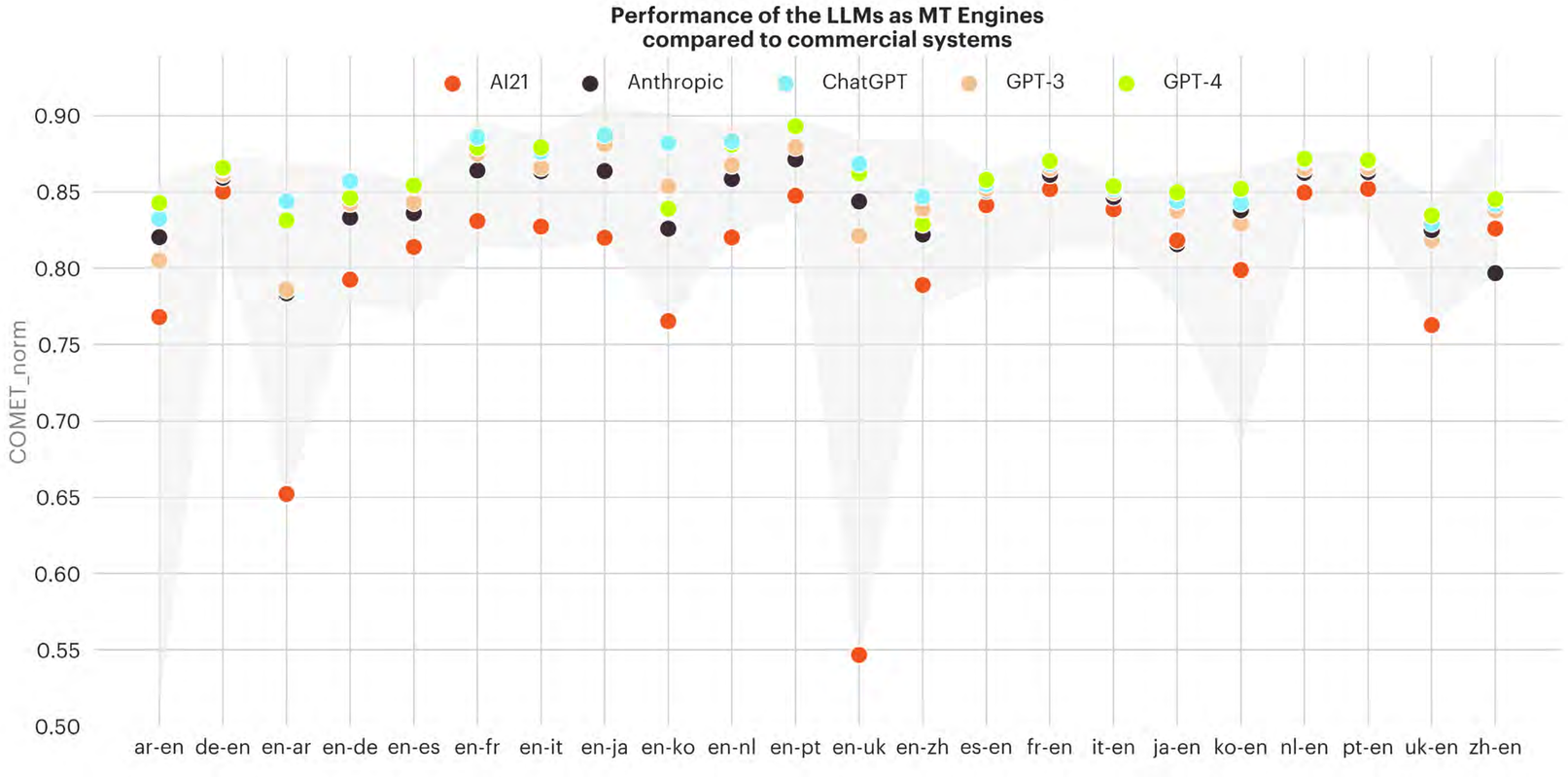
Claude is capable of a wide variety of conversational and text processing tasks, including machine translation, summarization, search, creative and collaborative writing, and Q&A.

Jurassic-2 by AI21

Jurassic-2 is the latest of the AI21's foundation models with such capabilities as zero-shot instruction-following, reduced latency, and multi-language support.

5.8 Large Language Models Performance (COMET)

- → Most models show great capability in translating to English: in most language pairs with English target, LLMs perform on par with classical MT commercial engines.
- → ChatGPT (gpt-3.5-turbo) and GPT-4 perform in the first tier compared to commercial engines in several language pairs with English source, such as en-pt, en-de, en-fr.
- → GPT-3 and Claude by Anthropic perform on par with the second tier of commercial engines.
- → Jurassic-2 by AI21 shows the lowest results out of all LLM models, most likely because the data it was trained on was skewed mostly towards English language.



6. Takeaways

6.1 Key Conclusions

1. The MT market is growing: 48 vendors

Three more vendors offer pre-trained MT models since the 2022 MT Report, plus there are several open-source pre-trained MT engines available. We have evaluated 37 MT engines — among them are 5 Large Language Models that have just come to the market.

2. MT covers 190K language pairs

190,085 unique language pairs across all MT engines. 65K more than last year and still growing. The main contributors are NiuTrans with their 112K, Alibaba with 16K, and NLLB by Meta with 5K unique language pairs.

3. 23 best performing MT Engines

23 MT engines are among the statistically significant leaders for 9 domains and 22 language pairs. 5 MT engines provide minimal coverage for all language pairs and domains, 2–3 per domain.

4. Large Language Models are in the 1st tier

Open-source MT from NVIDIA and TartuNLP mostly perform in the 2nd tier of commercial systems. Large Language Models, such as ChatGPT and GPT-4, perform on par with 1st tier commercial systems in most language pairs.

5. Two domains require a careful MT choice

Many engines perform best with English to Spanish, Portuguese, and Chinese, and in most pairs with English target. IT and Legal domains, as well as Arabic, Japanese and Ukrainian languages require a careful choice of MT vendor, as relatively few perform at the top level.

6. Two domains need customization

Despite having several comparable MT engines per language pair, Entertainment and Colloquial show relatively low scores, which may indicate the importance of customization in these domains.

6.2 Translate 20x More, With the Same Budget, Across the Company

The Machine Translation market is rapidly evolving — so it's crucial to re-evaluate the options to optimize localization budgets while maintaining the superior translation quality.

Evaluate and customize MT with your dataset on many platforms at once with Intento MT Studio or ask our experts for professional help.

Talk to experts

4,350

Evaluated stock and custom models by Intento

190k+

Language pairs available for evaluation

6.3 MT Evaluation & MT Maintenance for hassle-free Enterprise MT

MT Evaluation

- → Data cleaning
- → Model training
- → Test sample translations
- → Model training analysis
- → LQA (sample review)
- → Final analysis

Learn how to build or improve your MT program

MT Maintenance

- → MT Performance Monitoring & Hot-Swap
- → Glossary updates
- → Model updates
- → MT Quality Monitoring
- → Post-editing Effort Analysis
- → MT Evaluation

Learn how to evolve your MT program over time

Fast and Safe

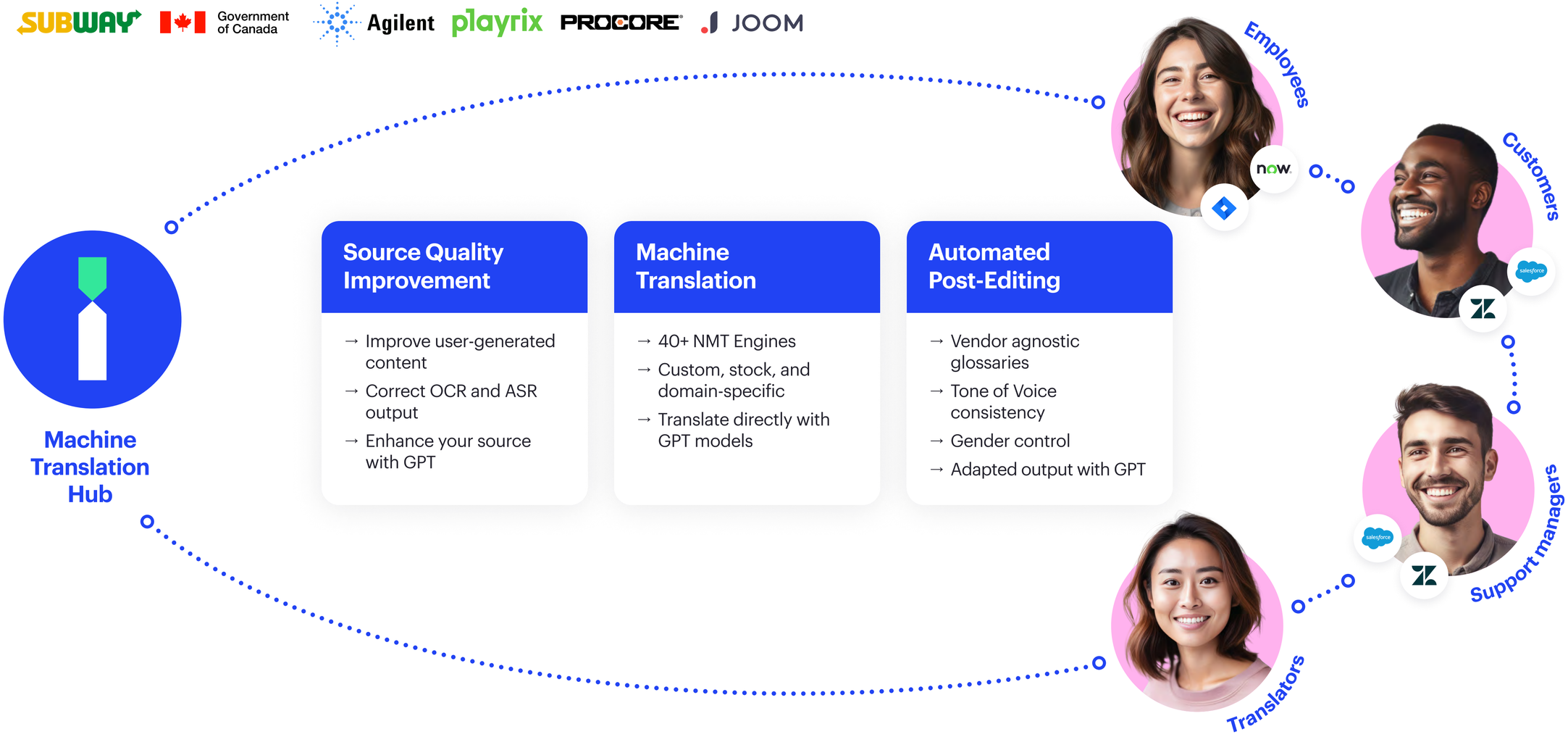
Only 5-6 weeks to get a winning MT engine with estimations for effort saved in post-editing and quality in real-time cases, such as support chats

Trusted

We run 15–20 MT Evaluation projects per month for global companies across industries under strict Security, Quality, and Data Protection requirements. ISO 27001 and ISO 9001 certified.

6.4 Enterprise MT Hub – Translation Backbone for Global Companies

SUBWAY · Government of Canada · Agilent · playrix · PROCORE · JOOM



6.4 Bring Machine Translation Benefits to Your Entire Enterprise With All-In-One, Scalable Solution

Unlock the power of machine translation and AI

Source Quality Improvement, Automatic Post-Editing, and quality estimation, normally done by humans, now managed via powerful algorithms.

Maximize efficiency and optimize headcount

Real-time translation is 200x more economical than human translation, with a 98% user acceptance rate.

Keep existing workflows running smoothly

Machine Translation Hub integrates with the most popular platforms, including human workflows in localization and document management.

Connect to 40+ machine translation providers with a single contract

One integration with Machine Translation Hub guarantees the best possible translation for your languages and domains.

Book a demo

Free

6.5 Master MT With Hands-On Guidance From Machine Translation University

Learn all about modern Machine Translation with our free, quick courses for localization, customer support, and employee experience pros.

→ **Why Machine Translation (MT): An introduction**
Solve your most pressing business problems through AI

→ **Selecting and Customizing Machine Translation Engines**
Leverage the entire MT landscape to your benefit

→ **Considerations for Selecting Machine Translation Engines**
Infrastructure and workflows

→ **Best Practices for Training MT Models using Translation Memories and Glossaries**
Preparation of data assets

Head to courses

The State of Machine Translation

An independent multi-domain evaluation of MT engines

Commercially available pre-trained MT models

2261 Market St, #4273
San Francisco, CA 94114

inten.to

3655 Nobel Drive, Suite 520
San Diego, CA 92122

e2f.com

Appendix A

A.1 Comparing hLEPOR and BERTScore

low hLEPOR + high BERTScore

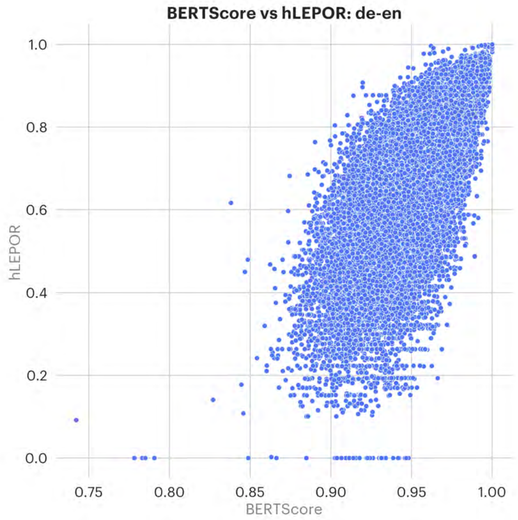
→ paraphrases / synonyms

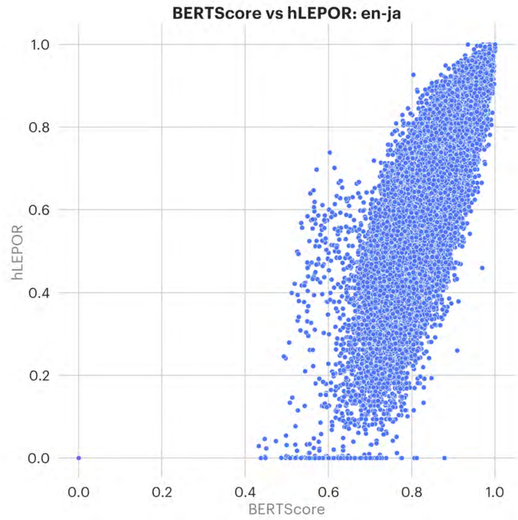
→ minor differences in plurality between reference and MT

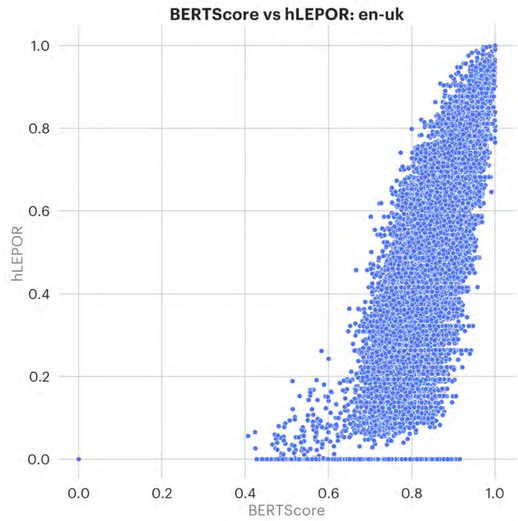
high hLEPOR + low BERTScore

→ mostly doesn't exist

→ punctuation and spacing issues







A.2 Comparing hLEPOR and COMET

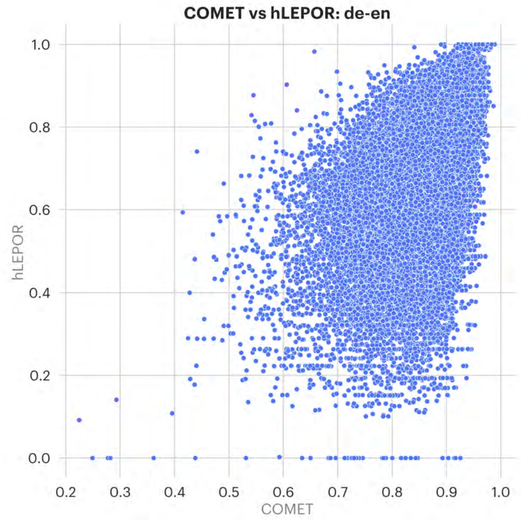
low hLEPOR + high COMET

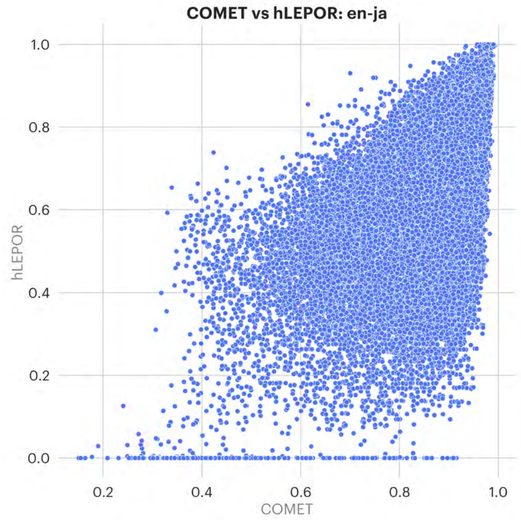
→ paraphrases / synonyms

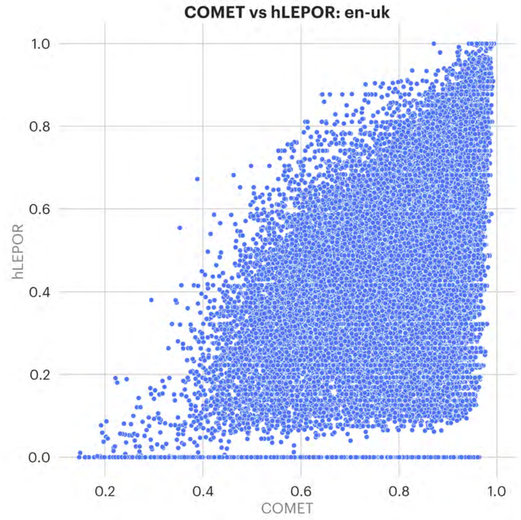
→ minor punctuation / tokenization issues

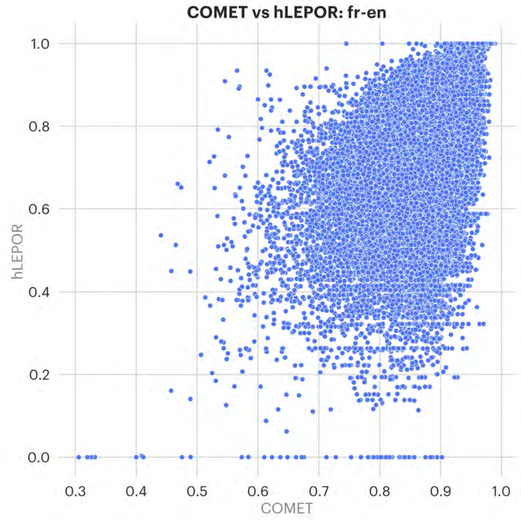
high hLEPOR + low COMET

→ COMET penilizes one-word omissions that do not affect hLEPOR that much









A.3 Comparing BERTScore and COMET

low BERTScore + high COMET

→ context-dependent alternative translations with different meanings (non-paraphrases)

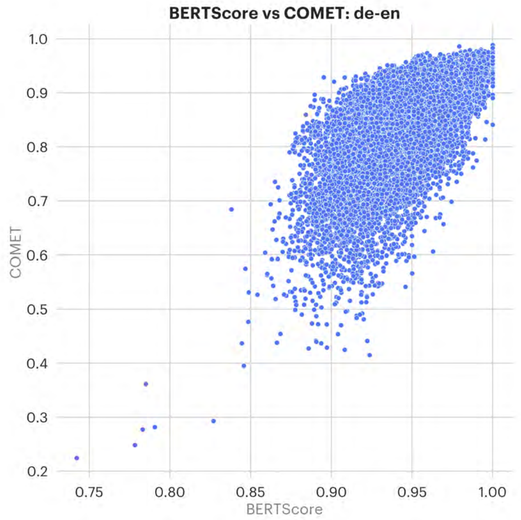
→ minor tokenization issues (e.g. merging words vs using “-“ in German)

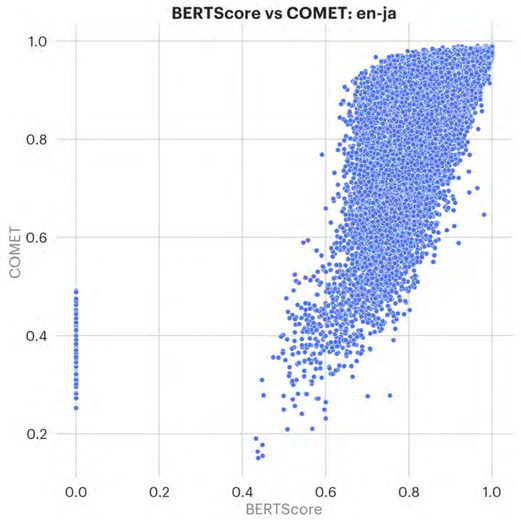
high BERTScore + low COMET

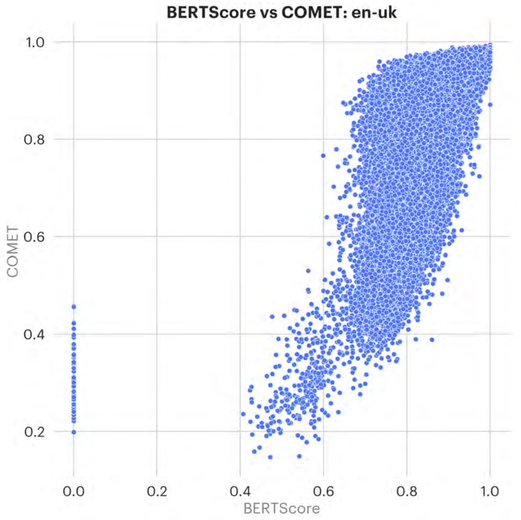
→ omissions and omissive paraphrases

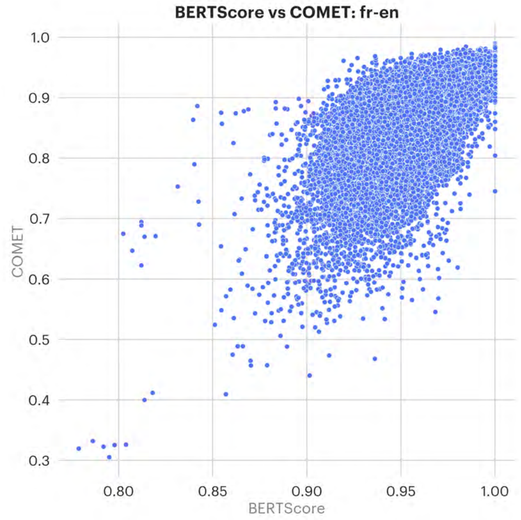
→ context-dependent alternative translations with a different gender or tone of voice (mostly short sentences that lack context)

→ difference in capitalization









Appendix B

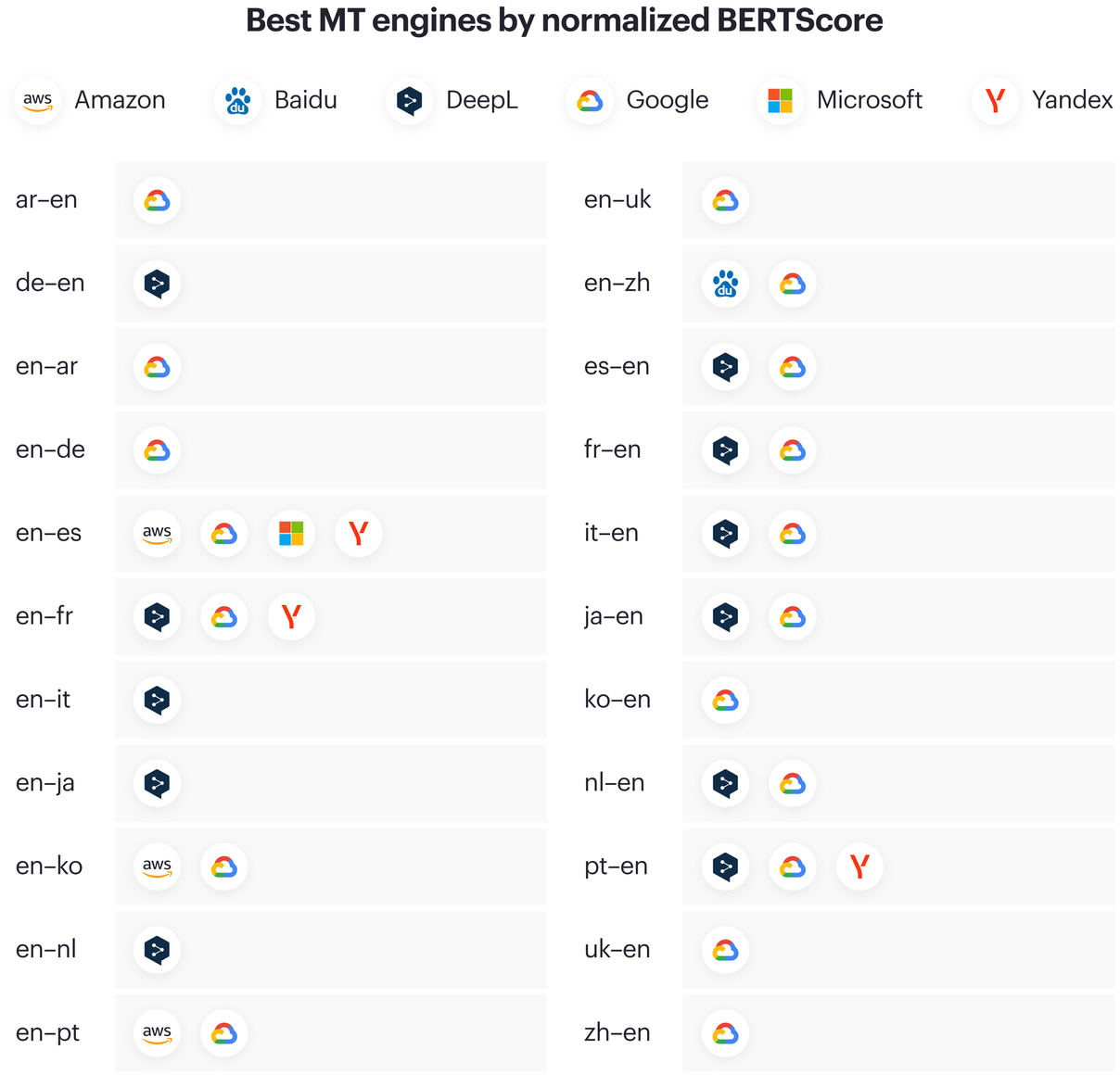
B.1 Ranking for BERTScore

- → For every language pair, we have normalized BERTScore to fit the [0,1] interval.
- → BERTScore rarely penalizes omissions and omissive paraphrases.
- → BERTScore penalizes different capitalization, therefore we have lowercased all text inputs. Per our observations, it does not lead to score corruption for properly capitalized sentences.
- → Does not reflect absolute quality level. Not comparable across language pairs.

MT vendors in one bucket provide the best quality for this language pair and domain, with no statistically significant difference between them. They are presented in alphabetical order.

B.2 Best MT per Language Pair (BERTScore)

- → There are slightly more leading MT engines than COMET suggests, 6, with a similar amount of engines per language pair.
- → 7 engines provide minimal coverage: Google, DeepL, ChatGPT, Amazon, SYSTRAN, Microsoft, and Naver.
- → Absolute values are not shown to avoid confusion, as the scores are not comparable across language pairs.
- → The domain and content type mix is different for every language pair (see the next slide) and greatly influences this leaderboard.



Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

B.3 Best MT Engines per Domain (BERTScore)

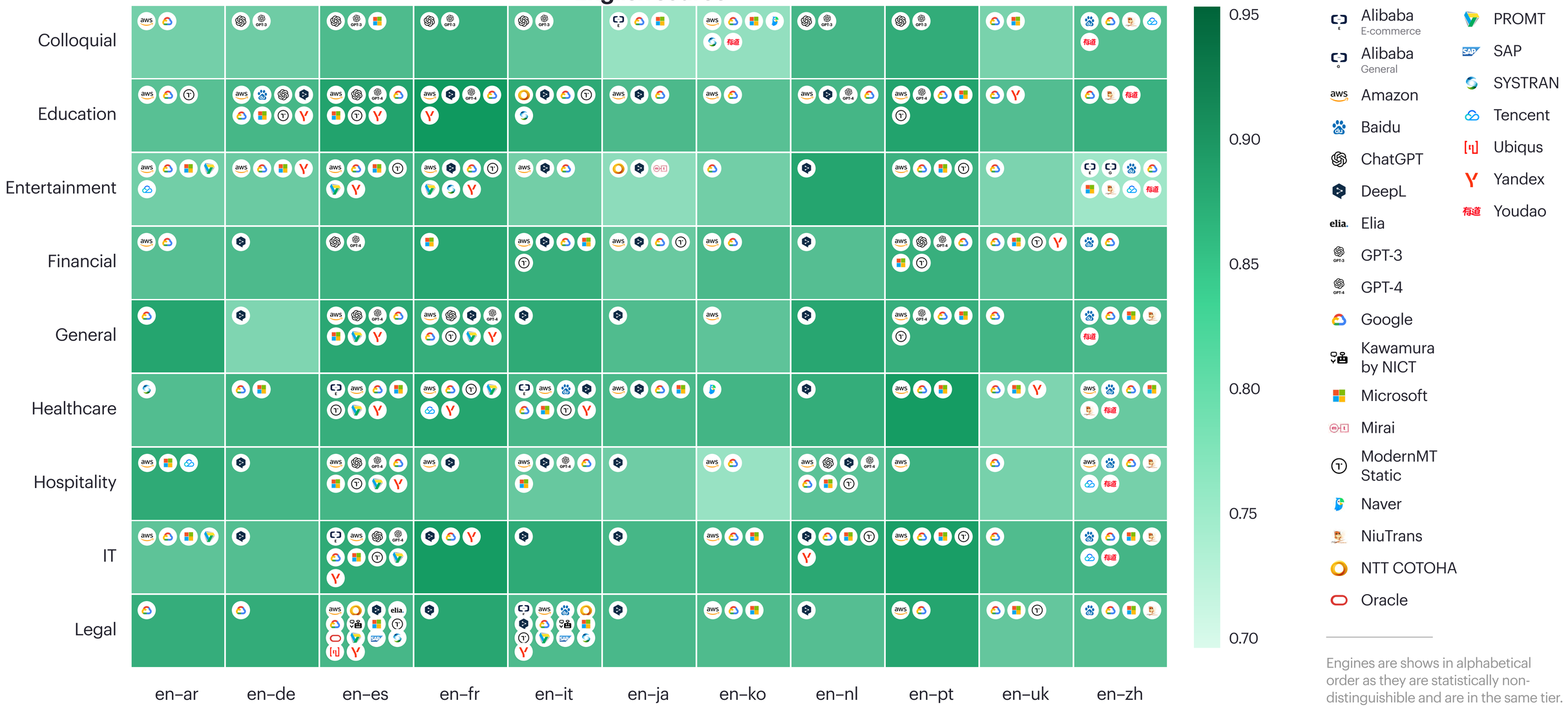
→ In the next slide, we show a heatmap of the best MT engines by a normalized BERTScore score. Each square shows the best providers for a particular language pair in a specific domain. The color of the square explains how high the best engines ranked among all engines in this combination of pair and domain.

→ For example, we see that the best engine for the English-French pair in the Education and Legal domains is DeepL. Its score for the Education domain is higher, and we expect less post-editing than in Legal.

→ Please remember that the absolute values of scores depend on the language pair you evaluate, and one should not compare scores between different language pairs.

MT vendors in one bucket provide the best quality for this language pair and domain, with no statistically significant difference between them. They are presented in alphabetical order.

Available quality and best commercial MT engines by domain per normalized BERTScore
English source



Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

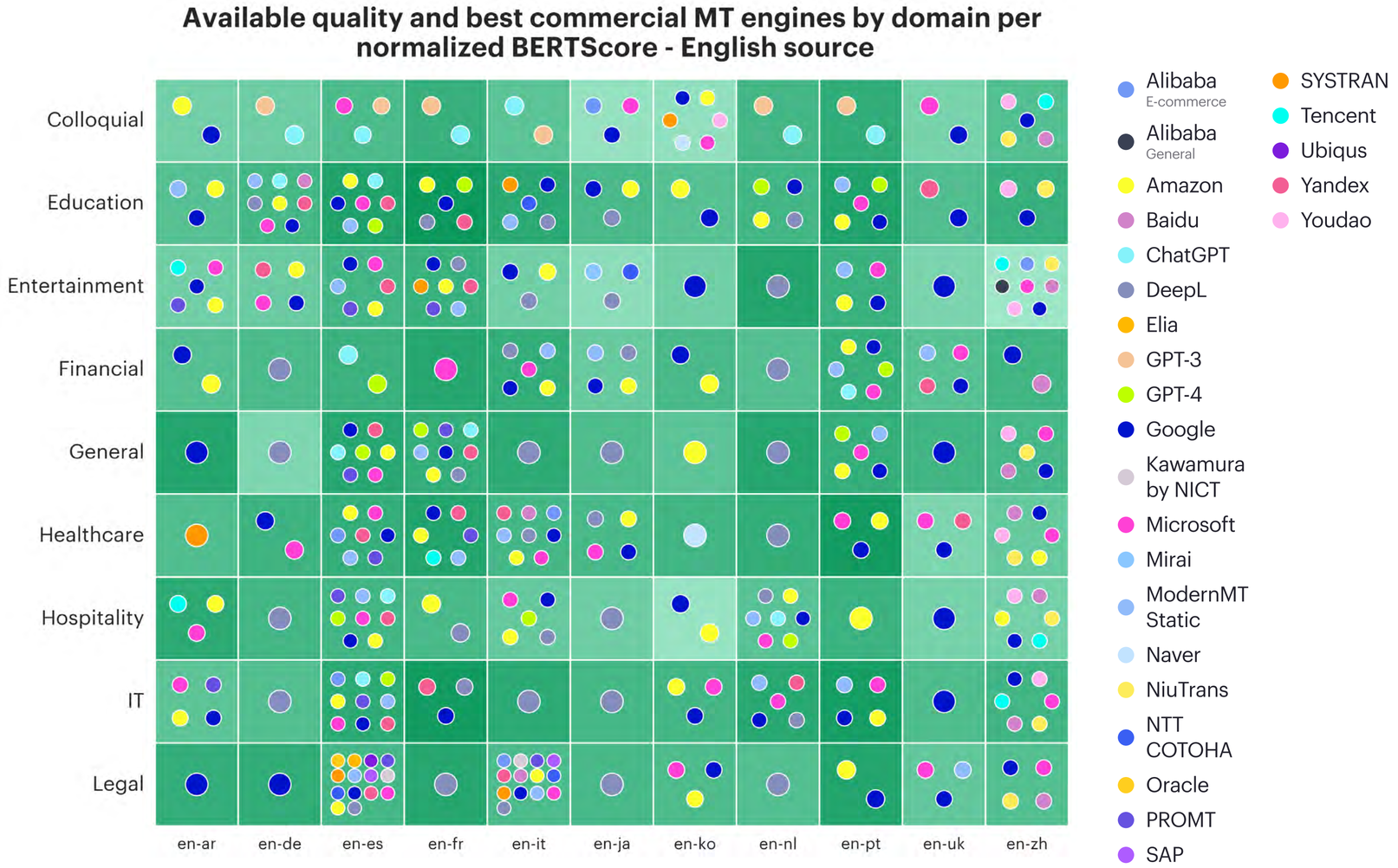
B.4 Best MT per Domain (BERTScore)

→ This chart is provided for reference. We recommend using the COMET chart on Slide 25 and Slide 27.

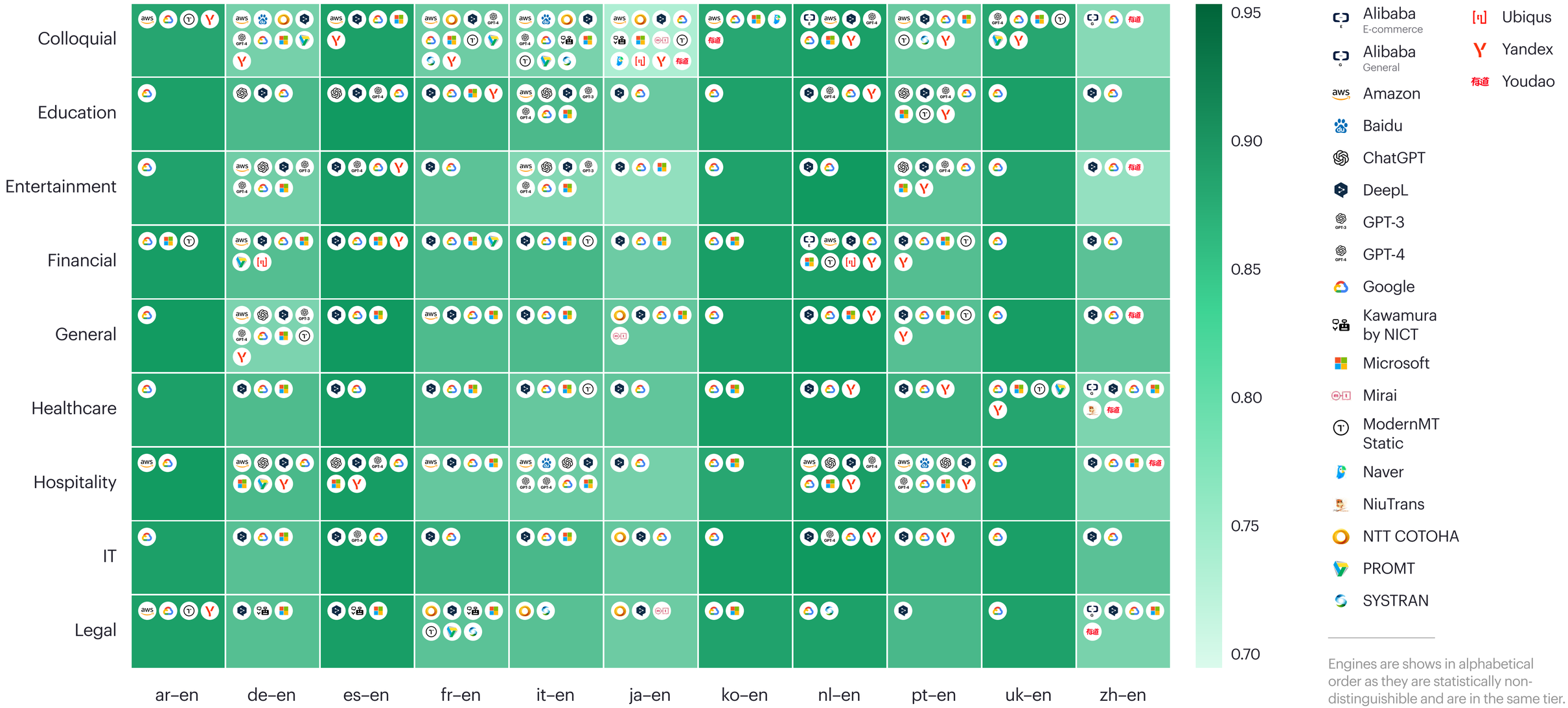
→ 25 MT engines are among the statistically significant leaders for 9 domains and 11 language pairs.

→ The only significant difference from COMET is English to Korean, Healthcare domain, where unlike COMET there is only one leading option, Naver.

→ BERTScore favors Google a lot – our hypothesis is that because BERTScore is a Google product it might be more familiar with its translation style.



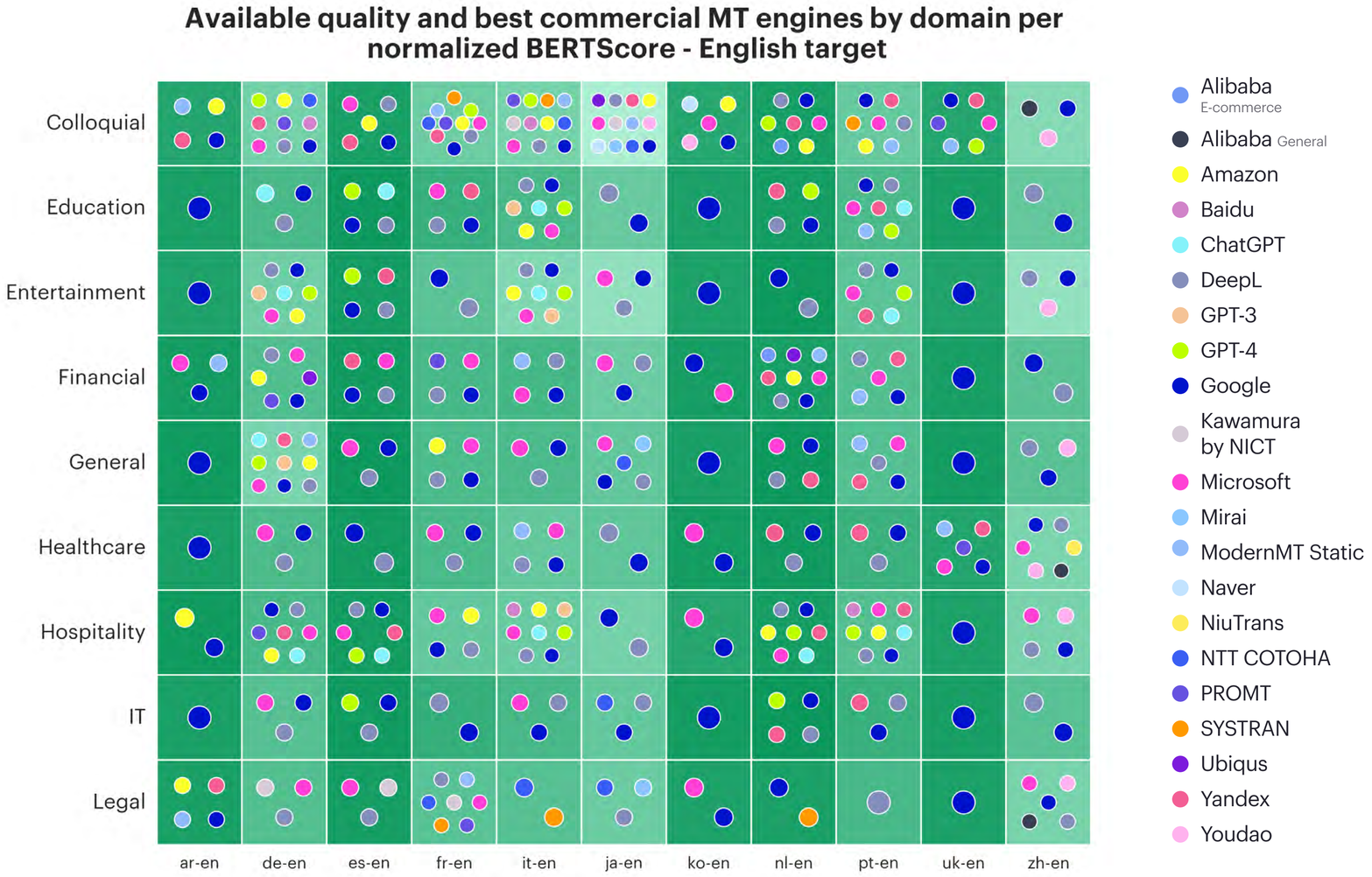
Available quality and best commercial MT engines by domain per normalized BERTScore - English target



Engines are shows in alphabetical order as they are statistically non-distinguishible and are in the same tier.

B.4 Best MT per Domain (BERTScore)

- → This chart is provided for reference. We recommend using the COMET chart on Slide 23.
- → 21 MT engines are among the statistically significant leaders for 9 domains and 11 language pairs.
- → Arabic to English, Ukrainian to English, and Korean to English require careful choice of provider as only a small numbers performs well.
- → BERTScore favors Google a lot – our hypothesis is that because BERTScore is a Google product it might be more familiar with its translation style.

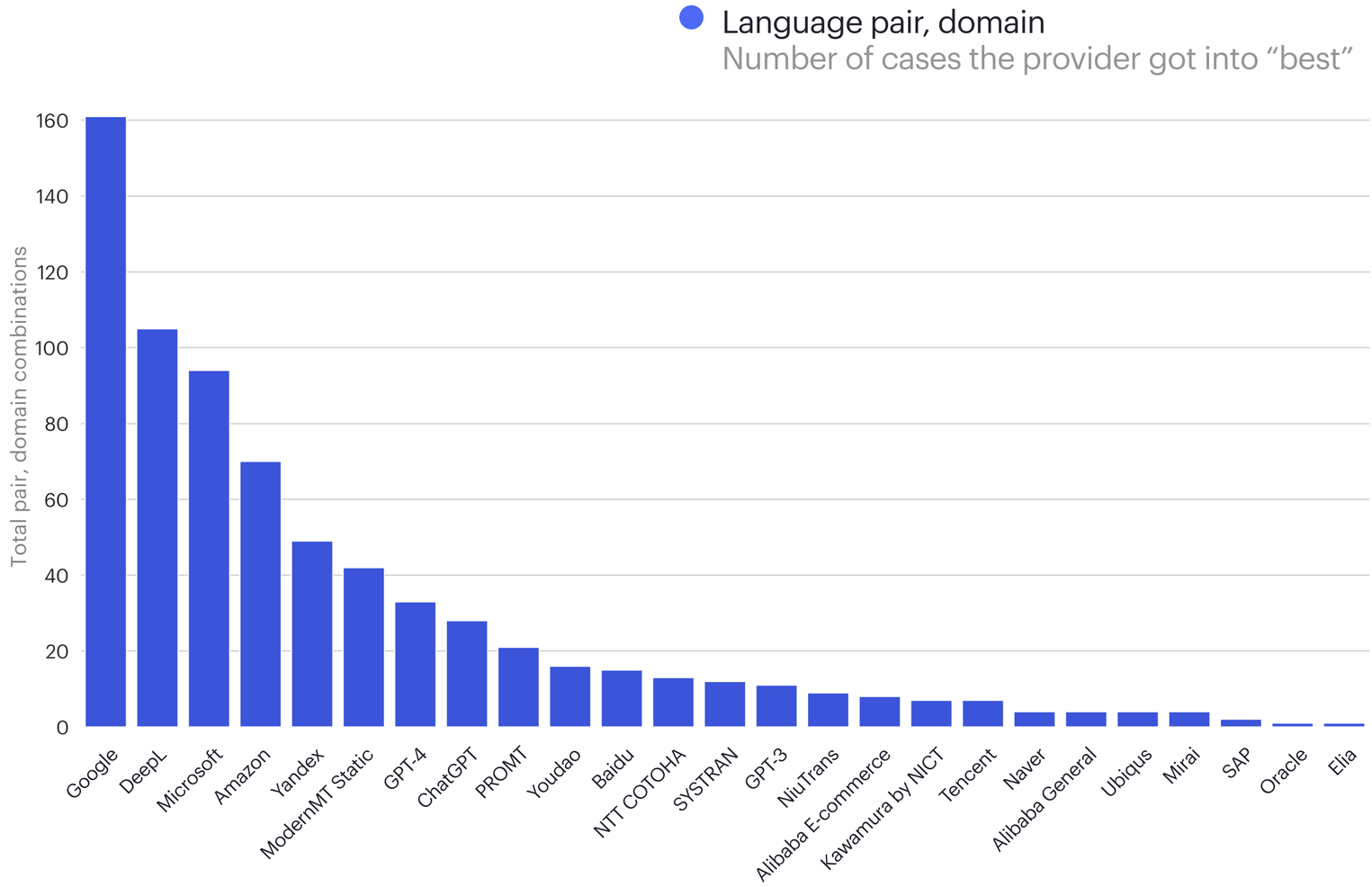


B.5 Top Performing MT Providers (BERTScore)

22 language pairs, 9 domains

Some providers were tested only in their specific domains and language pairs:

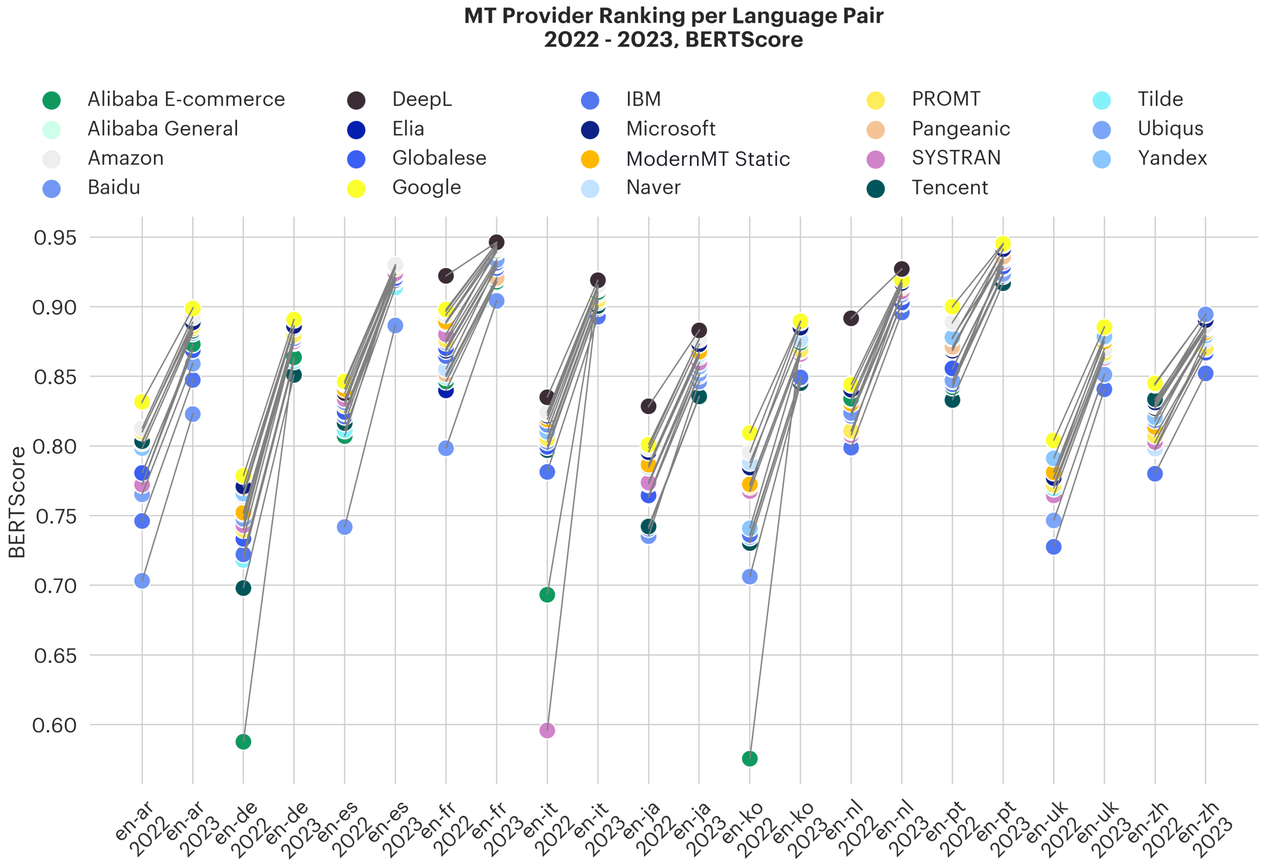
- → HiThink RoyalFlush specializes in en-zh translation in the Finance domain
- → TREBE specializes in Iberian languages, and was used for en-es and es-en translation
- → SAP is intended to be used with “SAP-specific” tasks and is not to be compared with the genetic stock models
- → Out of different ModernMT models, the static model was evaluated



B.6 Historical results (BERTScore)

→ All providers show growth in BERTScore compared to the previous year.

→ Providers with biggest score improvements include Alibaba E-commerce (for pairs en-de, en-it, and en-ko), Systran (en-it), ModernMT Static (en-it) and Ubiqus (en-ar, en-es, en-ko).



Appendix C

C.1 Scores for Sentences of Different Lengths

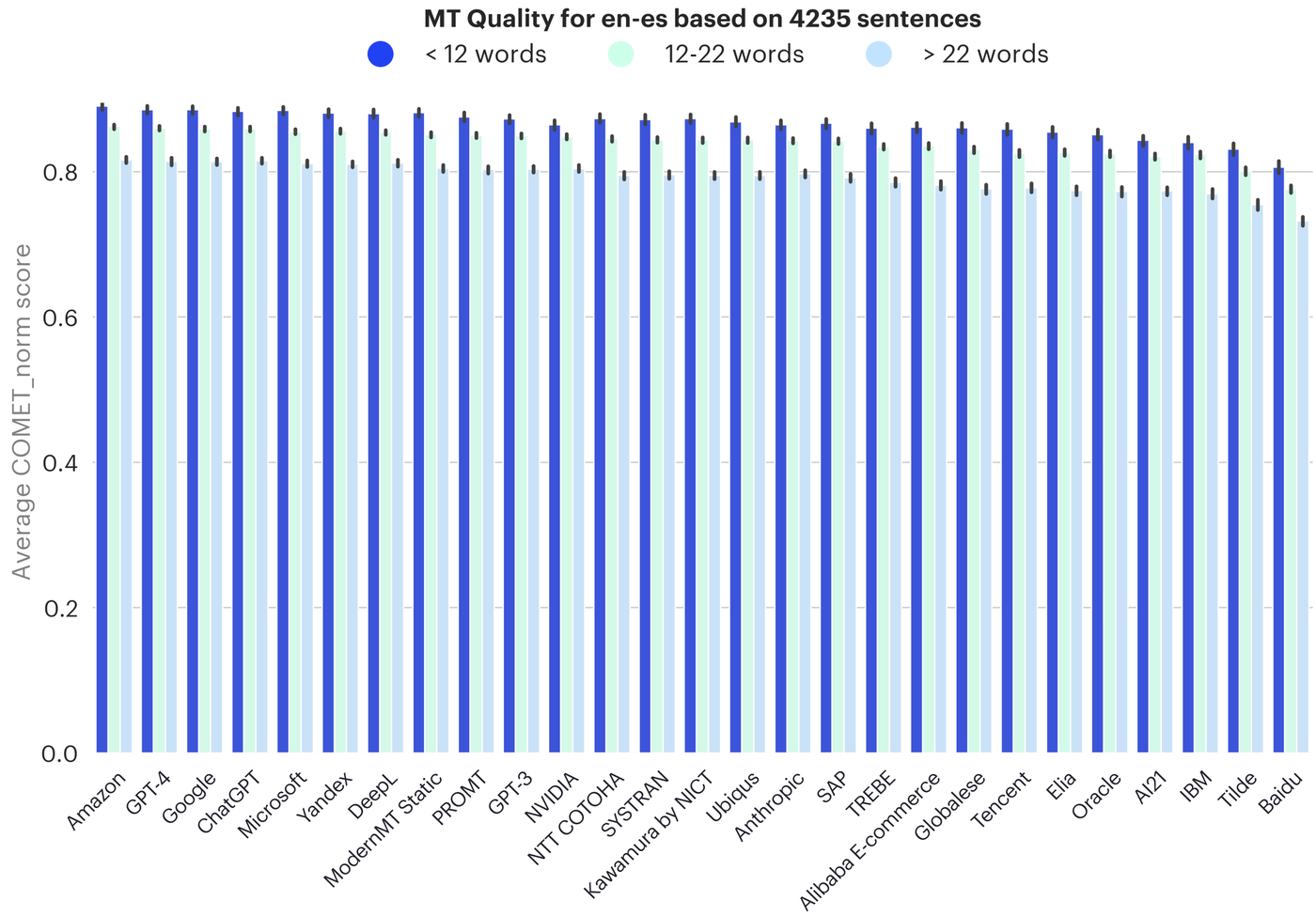
C.1 Scores for Sentences of Different Lengths

→ Typically, the scores are higher for shorter sentences.

→ English-to-Spanish demonstrates significant difference among MT engines for short and long segments (see the picture)

→ Some MT engines provide the top-tier scores for short and medium sentences, but fail to translate long ones, leading to the below average performance:

- GPT-4 for en-zh
- Amazon for en-uk
- IBM for en-de, en-uk



Appendix D

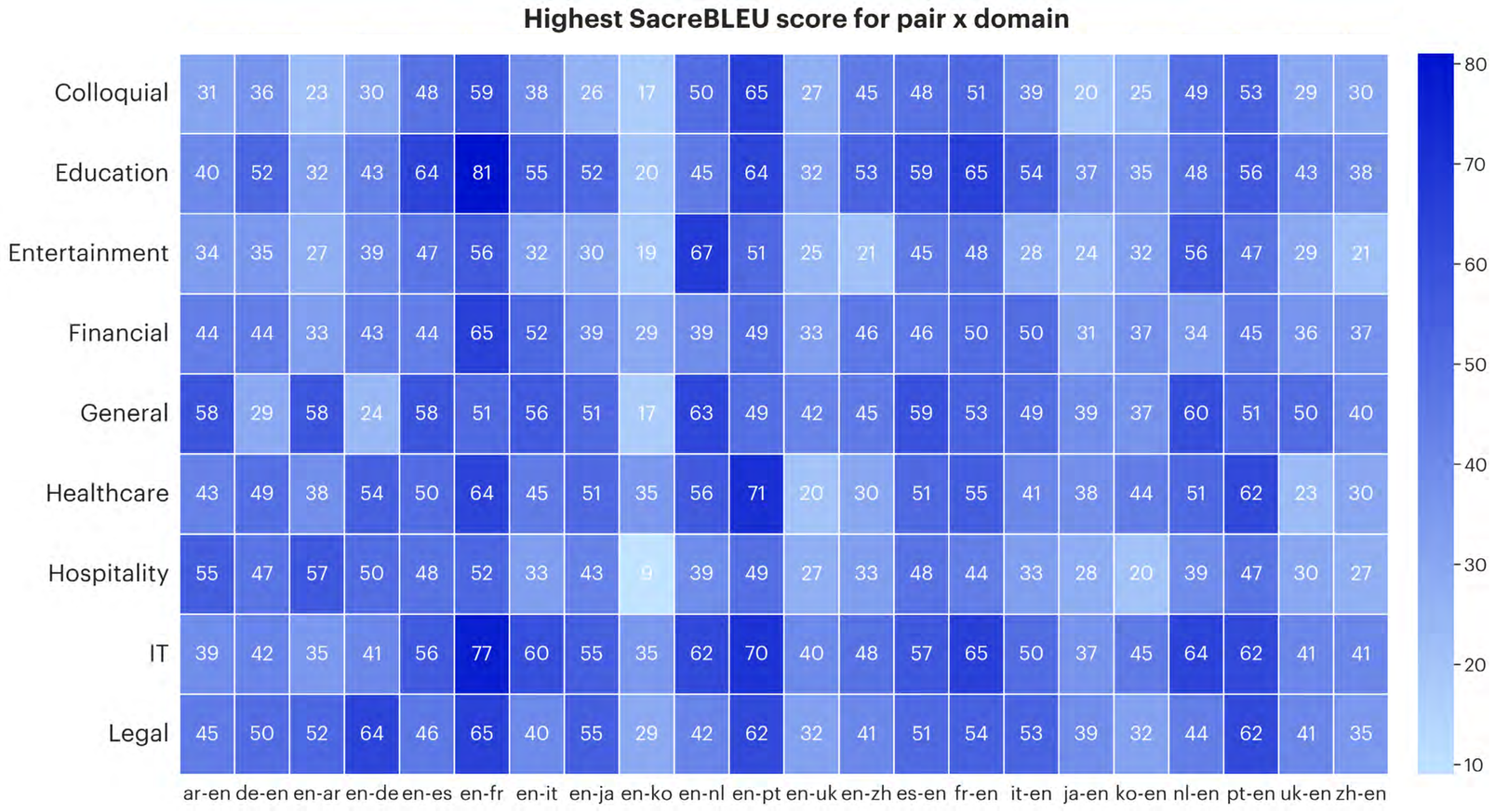
D.1 Best scores per Domain (BLEU)

D.1 Best scores per Domain (BLEU)

→ In the past, we were often asked “OK, but what are the BLEU scores”? Today, it’s commonly accepted that one should not use BLEU score at all. However, since you’ve asked for it, we decided to give you the highest SacreBLEU scores in each combination of domain and language pair.

→ There’s no statistical significance test as BLEU is a corpus-based score.

→ Please keep in mind that BLEU, as a corpus-level score with a number of parameters, is not comparable not only across different languages, but also across different datasets and different BLEU implementations.



Appendix E

E.1 2023 MT Leaders

E.1 2023 MT Leaders

MT Leader	A top-tier machine translation provider that ranked best in at least one pair or domain during the evaluation.
Language Leader	MT provider that is ranked best across the biggest number of domains within a language pair. The best rank in the maximum domains showcases their expertise and reputation as a leading MT provider for a specific language pair.
Domain Leader	MT provider that is ranked best in a specific domain within a biggest number of language pairs. The best rank for a specific domain signifies their expertise and reliability, establishing them as a trusted resource for delivering precise and high-quality translations within that particular domain.
Language Coverage	Language Coverage represents an extensive breadth of language support offered by MT provider. It serves as an indicator of the provider's capacity to handle 9,900 or more language pairs or cater to 100 or more distinct languages.

MT vendors in one bucket provide the best quality for this language pair and domain, with no statistically significant difference between them. They are presented in alphabetical order.

E.1 Alibaba

Alibaba Cloud

Online intelligent machine translation service developed by Alibaba Group. With the cutting-edge natural language processing technology and massive Internet data, Alibaba successfully launched its attention-based deep neural machine translation (NMT), helping users cross the language divide, share and access information smoothly, and achieve barrier-free communication.

Alibaba General ranked best in Colloquial, Healthcare, and Legal domains in Chinese (Simplified) - English language pair. Alibaba E-commerce ranked best in Colloquial domain in German - English language pair.

E.1 Amazon

amazon

Cloud-based neural machine translation service provided by Amazon Web Services (AWS) enables developers to easily incorporate language translation capabilities into their applications, products, or services. With its powerful machine learning algorithms, Amazon can accurately translate text between different languages, making it a valuable tool for businesses and organizations operating on a global scale. Whether it's translating product descriptions, customer reviews, or user-generated content, Amazon offers a scalable and reliable solution for multilingual communication.

Ranked best in 104 out of 198 pair/domain combinations. Language Leader in English - Arabic, English - Spanish (LA), English - Korean, and English - Portuguese (Brazilian) pairs.

E.1 Anthropic

ANTHROP\C

Anthropic is focused on understanding and building reliable, intelligent systems. The company employ a multidisciplinary approach, combining expertise in machine learning, computer science, and cognitive science to tackle complex challenges. Anthropic aims to develop AI technologies that are not only powerful but also aligned with human values and capable of augmenting human capabilities, leading to positive societal impact.

Ranked best in Entertainment domain in German - English, Spanish (LA) - English, and Italian - English language pairs. Also ranked best in Education in German - English, IT in Italian - English, and Hospitality in Portuguese (Brazilian) - English language pair.

E.1 Baidu

Baidu Translate is an advanced machine translation system developed by Baidu, a leading Chinese technology company. Powered by deep learning algorithms and neural networks, it provides accurate and fluent translations between multiple languages. The Baidu Translation API provides multilingual translation services through the HTTP interface. You only need to call the Baidu translation API, fill in the content to be translated, specify the source language (support the source language automatic detection) and the target language type to get the corresponding translation results.

Ranked best in 13 out of 198 pair/domain combinations.

E.1 Open AI

OpenAI

OpenAI is a renowned artificial intelligence research organization known for its cutting-edge advancements and contributions to the field. With a focus on developing and promoting friendly AI, OpenAI strives to ensure that artificial general intelligence benefits all of humanity. Through its groundbreaking language models, such as GPT-3, OpenAI has demonstrated impressive capabilities in natural language understanding and generation. OpenAI also promotes transparency and collaboration by releasing research papers and providing access to its models and tools, fostering innovation and pushing the boundaries of AI technology.

GPT-3 ranked best in 17 out of 198 pair/domain combinations.
ChatGPT ranked best in 61 out of 198 pair/domain combinations. Domain Leader in Entertainment.
GPT-4 ranked best in 75 out of 198 pair/domain combinations.

E.1 COTOHA

AI translation platform service "COTOHA® Translator" with AI engine that utilizes the latest "Neural Machine Translation (NMT) Technology" realizes over TOEIC 960 leveled high accuracy translations. Business documents such as Word, PowerPoint, Excel, and PDF can be translated in the same layout instantly, helping to accelerate your business.

Ranked best in Colloquial domain in Spanish (LA) - English, French (European) - English, Italian - English, Japanese - English, and Portuguese (Brazilian) - English language pairs. Also ranked best in Legal domain in Italian - English; General, Healthcare, and IT in Japanese - English.

E.1 DeepL

DeepL

Highly advanced machine translation system developed by DeepL GmbH. It utilizes deep learning algorithms and neural networks to provide accurate and fluent translations between multiple languages. With its exceptional translation quality and natural language processing capabilities, DeepL MT has gained recognition as one of the leading machine translation solutions in the industry, serving individuals and businesses alike in overcoming language barriers and facilitating effective communication across global contexts.

Ranked best in 126 out of 198 pair/domain combinations. Domain Leader in Legal. Language Leader in 13 pairs.

E.1 Google

Google

Advanced machine translation service designed to facilitate seamless and accurate translations across multiple languages. With its neural machine translation models, users can achieve exceptional fluency and precision in their translations. Google Cloud Translation AI supports a wide range of languages, including rare and low-resource ones, ensuring accessibility for global users. It also offers customization options through domain-specific glossaries and training of custom models, enabling users to tailor translations to specific industries or domains. Integration with other Google Cloud services allows for easy incorporation of translation capabilities into applications and workflows, supported by a robust and scalable infrastructure.

Ranked best in 158 out of 198 pair/domain combinations. Language Leader in 10 pairs. Domain Leader in Colloquial, Education, Entertainment, Financial, General, Healthcare, Hospitality, and IT.

E.1 Kawamura by NICT

A made-in-Japan machine translation engine powered by Japan's National Institute of Information and Communications Technology (NICT) that utilizes Japan's Translation Bank® framework. It features industry-specific engines, which are Generic, Patent, Finance (basic), Finance (investor relations), Legal, and Science (life science and medical field). Available as add-ins for Microsoft Word, Excel, PowerPoint, and Outlook and even on Slack with the Kawamura NMT Slack app. More than a million users communicate smoothly every day.

Ranked best in Legal domain in English - Spanish (LA), Spanish (LA) - English, French (European) - English, and Italian - English language pairs.

E.1 Microsoft

Microsoft

Microsoft Translator Text API, part of the Microsoft Cognitive Services API collection, is a cloud-based machine translation service supporting multiple languages that reach more than 95% of world's gross domestic product (GDP). Translator can be used to build applications, websites, tools, or any solution requiring multilanguage support.

Ranked best in 84 out of 198 pair/domain combinations.

E.1 Mirai

Advanced AI-powered translation solution designed specifically for corporate document translation and multilingual projects. Equipped with a neural machine translation (NMT) engine, it significantly improves translation accuracy, particularly for business communication and economic news. The user dictionary function allows for the translation of text and files in groups, streamlining the translation process. This AI-driven automated translation solution is highly favored by large corporations, as it supports multiple languages and a wide variety of file formats, resulting in enhanced productivity and efficiency in translation tasks.

Ranked best in Colloquial, Entertainment, General, Healthcare, and IT domains in Japanese - English language pair.

E.1 ModernMT

modernMT

Innovative machine translation system that draws inspiration from the human brain's adaptability and ability to learn from new experiences. It adopts a unique human-in-the-loop approach, where the machine and the human collaboratively assist each other throughout the translation process. This innovative methodology not only ensures outstanding translation quality but also enhances process efficiency, surpassing previous approaches in the field of machine translation.

ModernMT Static ranked best in 33 out of 198 pair/domain combinations.

E.1 Naver

NAVER

Leveraging its expertise in multilingual language processing, Naver applies advanced techniques to its translation engine, resulting in more accurate translation results. In contrast to technologies that solely rely on simple statistics, Naver actively incorporates user feedback to continually improve the quality of translations, addressing the limitations of traditional approaches. With a focus on providing reliable and precise translations, Naver's translation services offer enhanced language communication and understanding for its users.

Ranked best in Entertainment, Healthcare, and IT domains in English - Korean language pair. Also ranked best in Colloquial domain in Japanese - English and Korean - English language pairs.

E.1 NiuTrans

With its extensive language support, NiuTrans enables translation between any two of 304 languages. In addition to its Online Cloud service, NiuTrans offers on-premise deployment options and custom machine translation development services. The Xiaoniu Translation API provided by NiuTrans equips developers with seven translation service interfaces, automatically identifying source languages and handling XML tags and formats in texts. It seamlessly integrates with the Xiaoniu Translation Cloud Platform, granting direct internet access and unlocking the machine translation capabilities for websites, apps, and smart devices.

Ranked best in Colloquial, Education, General, Healthcare, Hospitality, and Legal domains in English - Chinese (Simplified) language pair. Also ranked best in Colloquial and Healthcare domains in Chinese (Simplified) - English language pair.

E.1 PROMT

PROMT®

PROMT is a leading provider of online translation technology, offering the PROMT Cloud API as a versatile cloud interface. This API enables seamless integration of PROMT's powerful translation capabilities into diverse programs and websites. With PROMT solutions, users can create applications, develop e-commerce platforms, and participate in international projects, all while benefiting from reliable and efficient translation services. PROMT's technology ensures accurate and timely translations, allowing users to navigate the constant flow of information with ease and facilitate effective communication across languages.

Ranked best in 16 out of 198 pair/domain combinations.

E.1 SYSTRAN

Offering a vast catalog of more than 50 languages, SYSTRAN has cultivated robust linguistic expertise to ensure accurate and reliable translations.Through collaboration with clients across diverse professional fields, such as legal, medical, and technical, SYSTRAN has developed tailored translation solutions that cater specifically to the unique requirements of each industry.

Ranked best in Colloquial domain in Spanish (LA) - English, French (European) - English, and Portuguese (Brazilian) - English language pairs. Also ranked best in Legal domain in English - Italian, Italian - English, Dutch - English, and Healthcare domain in English - Arabic language pair.

E.1 Tencent

Tencent Cloud

Tencent Machine Translation combines the advantages of neural machine translation and statistical machine translation. It automatically learns translation knowledge from a large-scale bilingual corpus and realizes automatic translation from the source language text to the target language text. At present, it can support the mutual translation of more than ten languages.

Ranked best in Colloquial, Education, General, and Hospitality domains in English - Chinese (Simplified) language pair.

E.1 Ubiqus

ubiqus

Ubiqus offers specialized translation engines tailored to meet the specific requirements of various sectors including law, medicine, finance, and more. Their engines are meticulously trained to address the specific requirements of each sector, ensuring accurate and tailored translations. By considering the context of the document, Ubiqus engines deliver grammatically correct translations that maintain the intended meaning and tone. Furthermore, clients have the flexibility to personalize their translations by incorporating their own terminology preferences, allowing for a seamless integration of industry-specific language in the translated content.

Ranked best in Colloquial and Financial domains in German - English language pair. Also ranked best in Financial and Legal domains in Dutch - English language pair.

E.1 Yandex

Yandex

With support for over 90 languages, Yandex Translate enables the translation of individual words or entire texts. Its robust functionality empowers developers to incorporate seamless and accurate translation services into their applications, facilitating effective multilingual communication across various platforms and industries.

Ranked best in 52 out of 198 pair/domain combinations. Language Leader in English - Ukranian pair.

E.1 Youdao

有道 youdao

Youdao is a leading provider of real-time, accurate, and convenient translation services for multiple natural languages. With its advanced technology, Youdao automatically collects vast amounts of multi-source parallel corpora, utilizing multi-machine, multi-card, and multi-lingual training techniques. This enables the development of optimized translation models tailored to specific domains, resulting in high-quality translations. Youdao's comprehensive range of services includes offline and online search, text translation, speech translation, and translation reading, providing users with versatile translation solutions to meet their diverse needs.

Ranked best in 14 out of 198 pair/domain combinations.